CCNA/CCNP课程笔记

主讲：李勇
撰稿：胡军
整理：熊伟
资料提供：陈翔

2004 年 10 月
VER 1.0
WWW.SOEASY.NAME

传输介质：

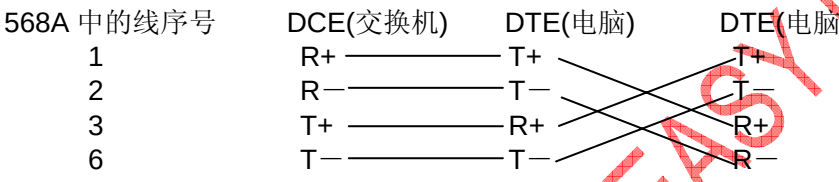
- 非屏蔽布线：西蒙公司、朗讯公司
- 屏蔽布线：IBM
- 光纤：不受电磁干扰、并且距离长
  多模光纤：内芯 6.25 微米，外芯 120 微米。
  单模光纤：内心 9 微米
- 无线：频率 2.4GHz(802.11b/g　100 米内),5GHz（802.11a　55 内）短波，衍射小、怕障碍物。

流行的传输介质：

名称	速率(bps)	有效距离(M)	线缆描述	对应标准
100base-T	100	100	4 芯（3 类以上双绞线）	IEEE802.3u
1000base-T	1000	100	8 芯（5 类以上双绞线）	IEEE802.3ab
1000base-CS	1000	25	25 欧姆屏蔽双绞线	
1000base-SX	1000	500	多模光纤、短波	IEEE802.3z
1000base-LX	1000	550	多模光纤、长波	
		3000	单模光纤、长波	

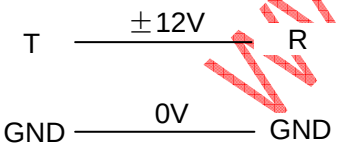
光纤接头：ST:跳线柜上用，很牢固。
SC:设备上用。

双绞线的线序：

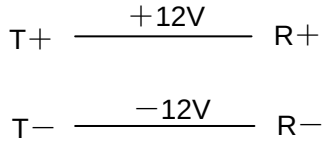


线序标准：
568A: 白橙 橙　白绿 蓝　白蓝 绿　白棕 棕
568B: 白绿 绿　白橙 蓝　白蓝 橙　白棕 棕

信号的发送方式：
平衡发送：



差分发送：抗干扰强，差值 24V 为 1，差值 0V 为 0（一对绞线）



RAM:running-config，正在运行的 IOS 系统
NVRAM:startup-config
ROM:IOS 最原始的、核心的部分
FLASH:IOS 最新的、最完整的版本

MAC（media access control）应用方式有 2 种：
争用 CSMA/CD(Carrier sense multiple access/collision detect 载波监听多路访问/冲突检测) 使用 CSMA/CD

介质访问模式的例子有 Ethernet、IEEE 802.3 网络

共享：Token-ring

MAC 是适合在子网内的操作。

局域网帧的几种标准：

1. 以太 II 帧（ethernet II）：

Flg	Smac	Dmac	Potocol type	Data	Fcs	Flg
起始符	源 mac	目的 mac	协议类型 8080：IPX 1XXXX：IP	IP 包或者 IPX 包等，有协议类型决定的	帧检测序列（校验和）	结束符

2. IEEE802.3 帧：

   与以太 II 帧的区别是，协议类型变为了长度。所以它不支持上层（网络层）多路复用，只能一种网络层类型，IP 包或 IPX 包。

3. IEEE802.3－SAP 帧：

   在 IEEE802.3 帧的 DATA 字段中插入 IEEE802.2 的 SAP 头

OSI 模型	IEEE802．2SAP 头	IEEE802.3 帧	以太 II 帧
LLC 逻辑链路层 MAC	LLC	 MAC	LLC MAC
物理层		物理层	物理层

4. IEEE802.3-SNAP 帧：

   在 SAP 头中再加入以太 II 帧的**协议类型**指示

信号发送的方式：

1. 单播：发给特定一个主机的动作
2. 广播：发给特定所有主机的动作

   以太网广播帧的目的 MAC: FF: FF: FF: FF: FF: FF
3. 组播：发给特定部分主机的动作

概念：

冲突域：共享介质的总线型主机群。（它是引起以太网性能下降的原因）

HUB 内部仍然是总线型，只是检测故障方便了。

交换机：1、分割冲突域 2、互联冲突域

交换机每个接口都是一个单独的冲突域。(microsegnention 微分段技术)

交换机≥HUB+网桥（交换机前身是网桥，网桥是软件实现的所以慢，端口也少。）

MAC address tab:记录端口与冲突域的主机对应关系。

桥接方式：（以太网交换机桥接是透明桥接）

1. 透明桥接：应用 MAC address table 转发帧，2 层操作。（不改变帧，所以叫透明交接）
2. 源桥接：改变帧来实现的 2 层操作。

交换机功能：（含透明桥接的功能）

1. 转发，是并行的
2. 过滤的转发：在有需要的接口转发。
3. learning:学习生成 MAC address table(条目生存时间 ARPtime 是 300 秒)
4. 广播，从所有接口转发帧。有两种帧会被转发
   a) 广播帧。
   b) 在 MAC address table 中没有匹配条目的帧。

路由器的一个接口就是一个广播域也是一个冲突域（大的广播域性能下降，所以要分割广播域。但路由器的价格高！）

VLAN:是一组交换机接口的集合，可跨越交换机。（一个 VLAN＝广播域＝一个子网）

Trunk link:可载波多个 VLAN 的线路。

Trunk port:连接多个 VLAN 的接口。

Access port:从属一个 VLAN 的接口，只能转发和接收这个 VLAN 的帧。

Tag 技术：帧标记技术,在同一个交换机上，用来区别该帧属于哪个 VLAN,

1. ISL 帧标记（cisco 专用）:
2. IEEE802.1Q 帧标记(国际标准)，或叫 dot 1q:在帧中插入字段。

ICMP(internet 消息控制协议)是 TCP/IP 的辅助协议。作用：用来通报错误的

Echo-request　echo 请求

Echo-reply　　echo 应答

命令 ping 就是调用 ICMP 中的

ARP（地址解析协议）是 TCP/IP 子协议

作用：已知目标的 IP，求其 MAC 地址。

过程：发出以太网广播帧，内容是请求已知 IP 的对应 MAC 地址.ARP 回应是单播帧。

路由过程：路由器解帧，提出其 IP，比较路由表（用其 IP“与”每个条目掩码，再看与哪个网段匹配），找出发送接口，封装帧，发出。（此过程中 IP 包的内容始终不变）

routed protocol（被路由协议，处于 3 层）:提供信息。如 IP 协议

routing protocol（路由控制协议，处于 7 层）：控制选择。如：RIP　IGRP　EIGRP　OSPF 协议

主类地址:X.X.X.X (或称主类边界)

A 类地址: X 为 1~126　　其主类地址的网络地址(或称主类边界)为 X.0.0.0（0xxxxxxx.0.0.0）

B 类地址: X 为 128~191　其主类地址的网络地址(或称主类边界)为 X.X.0.0（10xxxxxx.x.0.0）

C 类地址: X 为 192~223　其主类地址的网络地址(或称主类边界)为 X.X.X.0（110xxxxx.x.x.0）

D 类地址: X 为 224~239　其主类地址的网络地址(或称主类边界)为 X.X.X.0（1110xxxx.x.x.x）

E 类地址: X 为 240~248　其主类地址的网络地址(或称主类边界)为 X.X.X.0（11110xxx.x.x.x）

A 类地址的二进制表示:0　+ 7 位网络号　+ 24 位主机号

B 类地址的二进制表示:10　+ 14 位网络号　+ 16 位主机号

C 类地址的二进制表示:110 + 21 位网络号　+ 8 位主机号

保留地址:

10.x.x.x

172.16.X.X~172.31.X.X

192.168.X.X

静态路由条目输入：

IP route　IP 地址　掩码　下一跳 IP　AD 管理距离

其中下一跳 IP：一般是邻居的直连接口，也可以是可到达（即已有到该接口的路由条目存在）的非邻居接口，也可以是自身的接口（网络拓扑必须是 point-to-point 类）

如：IP route 10.1.20.0 255.255.255.0　10.1.3.2

在路由表中就会出现：　S　10.1.2.0/24　10.1.3.2　s0(s0 是端口号)

缺省的管理距离：

直连	静态路由	Eigrp 汇总	EBGP	Eigrp D	Igrp I	Ospf O	IS-IS	Rip R	Ospf IA	Ospf E1	Ospf E2	Eigrp EX	IBGP	未知

路由C	S	条目												的
0	1	5	20	90	100	110	115	120				170	200	255

RIP 动态路由协议

原理:周期性的将完整路由表转发给邻居.

1. 路由更新周期:30 秒
2. 转发的完整路由表名字叫 update 包.是从路由器每个接口上以 IP 包广播出去的.(是个 3 层模型的 IP 广播,不是 2 层模型的帧广播,2 层是以 MAC 值为基础的,而这个是以 3 层模型的 IP 地址广播出去的)
3. 路由无效时间（Invalidation timer）：当一条更新的路由条目建立时，路由无效时间这个计时器初始化为 180 秒，并开始往下递减，如一直收不到该条目的更新信息，当该计时器递减为 0 时，该条目被设为无效（具体的操作是：跳数被设为 16，即不可达）
4. rip v1 与 rip v2 的区别在于：v1（是 classfull 路由协议）发送的条目不含掩码。V2（是 classless 路由协议）则含。

在 rip v1 中，因为其发送 update 中的条目不含掩码（因为它是 calssfull 的），为了保证掩码相同，所以在发送条目之前它要进行是否发送的判断。

1. 发送方的判断：

发送端口判断	判断过程		备注
是否属于水平分隔	是：不发	否：发	防止路由环路一个方法
要发的**条目**和发送**端口**的主类地址	主类地址相同：比较掩码	掩码相同：发送	子网划分的深度一样的就发送
		掩码不同：不发	不支持变长子网，因为 classfull
	主类地址不同：自动汇总到主类边界发送		

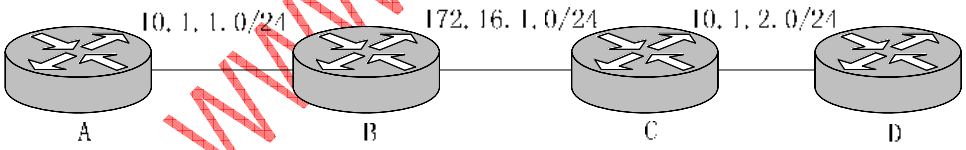
2. 接收方的判断及操作：

接收端口判断	判断过程	备注
收到的**条目**和接收**端口**的主类地址	相同：用接收端口的掩码填充该条目	不支持变长子网
	不同：用收到条目的主类地址掩码填充该条目掩码	

RIP 总结：

主类地址相同的网络地址划分出来的子网的掩码要一样长，不一样长的肯定会出错；

主类地址相同的网络地址划分出来的子网如被其它主类子网隔开了，肯定会出错。



B 的路由表：R 10.0.0.0/8 A 1hop R 10.0.0.0/8 C 1hop ，B 会认为是 A、C 都是可以到 10.0.0.0/8 的主类网段，且跳数一样，所以是并行路由，负载均衡自动执行。

所以，运行 RIP V1，主类地址相同的网络地址划分出来的子网掩码要一样长且不能被其它主类子网隔开。

命令格式:

```
router rip
  network  主类地址的网络地址
```

(该命令的含义:1.发出 update 包的接口有那些. 2.发包接口的 IP 地址被发送给邻居了)

rip 的 network 命令跟的不是主类地址的话，也会自动汇总为主类边界发送。

RIP V1 的缺点:（UDP 端口号:520）

1. 带宽占用高(周期性的 IP 广播完全路由表)
2. hop 跳数这个参数无法表示带宽
3. routing loop 路由环路：如果在某个广播周期的第 15 秒某个网段 down 了，自己就没了这个网段的条目。在

下个广播周期，邻居又向它发了这个网段的路由（每个路由器都是发它所有知道的网段条目），routing loop 出现了。

4. 收敛速度慢
5. 直径太小，最大 16 跳
6. rip v1 不支持 VLSM 可变长子网掩码
7. rip v1 不支持不连续的主网分配。

解决 Routing loop 路由环路的办法（cisco 用的前 4 个方法）

1 trigger update 触发发送：拓扑改变后，马上发出 update，而不用等到下个 update 周期。
2 定义了最大跳数 15，达到 16 则会产生无效。（该功能的实施是依靠 IP 包中的 TTL 字段）
3 holdtime=180 秒，如收到一个条目的跳数大于该条目原有的跳数，就保持该条目 180 秒不变（hold）.（cisco 在 180 秒之后又加了个 60 秒的 trash 时间，用来广播该网段的失效，加快其他路由器的收敛）
4 水平分隔：从路由器一个接口收到的路由条目不再从该接口发送回去。
5 水平分隔病毒反转：与水平分隔不同在于，还是回送，但发送的是 16 跳的条目（病毒反转）

执行rip 2的动态路由标准　version 2
同时要关闭自动汇总 no auto,因为默认是开的。

Igrp 与 rip 的不同：

- 直径：rip 最大 15，igrp 最大 99
- 能表示带宽。

访问列表:从一个集合中挑选出某些有特定标识的子集的工具.(只是挑出来,并没有执行什么)

访问列表作用:

1. 过滤IP包
2. 过滤路由条目
3. 增加跳数

命令格式:

标准访问列表:

access-list 1~99 permit /deny　源IP地址 网络掩码的反码（是匹配主机位用的)

扩展访问列表:

access-list 100~199 permit /deny 协议类型(IP /TCP) 源IP地址 匹配掩码 eq端口号 目的IP地址 匹配掩码 eq端口号

注意点:

1 标准访问列表中，当掩码是0 . 0 . 0 . 0时省略它，如果省略了掩码，则表示该掩码是0 . 0 . 0 . 0可以在IP地址前写h o s t来代替掩码0 . 0 . 0 . 0。
2 永远有一条deny any的语句在每个列表末尾,它是不显示出来的.ANY是0.0.0.0 255.255.255.255
3 in是先in后路由,out是先路由再out
4 在路由表的一个接口上,最多只能有一个in和out
5 看访问列表的命令是show access-list
6 掩码的反码只跟着访问列表和network（IGP类的，BGP的不是反码）后用，0表示必须匹配，表示了是多大的网络被包括进来了。1就是任意了，1表示的是包括的网络内的主机号码（1所在的位是主机位），一般都是要求匹配网络位，主机位任意。
而掩码是表示网络位的位数，还包括了子网络位的位数。

标准访问表尽量放在靠近目的端口的位置，以为它只有源的信息，无目的的信息。
扩展访问表尽量放在靠近过滤源的位置
目的：不会影响其它接口上的上的数据

访问列表配置要点：

- 访问列表的编号指明了使用何种协议的访问列表

- 每个端口、每个方向、每条协议只能对应于一条访问列表
- 访问列表的内容决定了数据的控制顺序
- 具有严格限制条件的语句应放在访问列表所有语句的最上面
- 在访问列表的最后有一条隐含声明：deny any－每一条正确的访问列表都至少应该有一条允许语句
- 先创建访问列表，然后应用到端口上
- 访问列表不能过滤由路由器自己产生的数据

标准访问列表和扩展访问列表比较

标准	扩展
基于源地址	基于源地址和目标地址
允许和拒绝完整的TCP/IP协议	指定TCP/IP的特定协议和端口号
编号范围 1 到 99	编号范围 100 到 199.

例子:

1 用访问列表,过滤IP包

access-list 1 deny 10.1.3.0 0.0.0.255 挑选

access-list 1 permit any

int e1

ip access-group 1 in 执行动作

2 用访问列表,过滤路由条目

rip eigrp bgp的in、out方向和ospf的in方向可用。（ospf交换的是LSA，不是路由条目。但在in的方向，SPF算法从LSA database推出路由表时起作用，之前无作用）

access-list 100 permit tcp 10.1.2.0 0.0.0.255 any eq 21 (这里的 any 是 0.0.0.0 255.255.255.255)

router rip

distribute-list 100 in e0 （仅对e0这个接口有效，对该路由器其他接口不限制。也可以不设“接口”这个参数，那将对所有接口都有效）

3用访问列表,增加跳数(update更新包中的路由条目的跳数)

access-list 2 permit 20.0.0.0 0.0.0.255

router rip

offset－list 2 in 3 e0 (3是加的跳数,最好在in的方向加跳数；仅对e0这个接口有效，对该路由器其他接口不限制。也可以不设“接口”这个参数，那将对所有接口都有效)

clear ip route * (马上收敛,可立刻用show看到)

路由器阻止更新包从某个接口发出

router rip

passive-interface lo0

passive-interface lo1

passive-interface lo2

动态路由协议分类:

按技术分:1 距离矢量类:rip igrp(cisco私有的) bgp

2 链接状态类:ospf is-is(CLNS迁移过来的，严格按7层模型来的，性能比ospf好，但用的少)

3 杂合类(包括以上两种的技术):eigrp(cisco私有的)

按有无类别路由选择协议分(classfull还是classless):1 classfull:rip v1 igrp

2 classless: eigrp bgp ospf ripv2

按从重分布类型分:1 IGP （内部网关协议）:eigrp bgp ospf rip v1 rip v2 is-is

2 EGP（边界网关协议）: bgp

自治系统AS:由共同的管理者管理的一大组路由器(比如中国电信与美国电信就是两个AS)

路由域:运行相同的路由协议,动态生成全网络拓扑图的一组路由器.(有时也被人称为AS,是不准确的)
自治系统AS可包含多个运行不同动态路由协议的路由域
运行不同动态路由协议的路由域之间以IGP重分布的方式连接
自治系统AS之间以EGP重分布的方式连接

概念:

1 子网划分:一种IP地址分配方案,划分主类地址位多个子网

例子:172.16.0.0/16只是一个B类的主类地址,掩码16位,即网络位还是16位,没有用VLSM可变长子网掩码.应用VLSM(可变长子网掩码)将网络位扩至24位,即有8个主机位变成了网络位,网络位多了256(2的8次方)个变化,所以这一个子网可以划分成256个子网.

而且已划分的子网还可以继续划分子网,如上述一个已划分的子网172.16.2.0/24再划分成2个子网,为172.16.2.0/25到172.16.2.127/25 和 172.16.2.128/25到172.16.2.255/25 两个子网.

即172.16.2.00000000到172.16.2.01111111

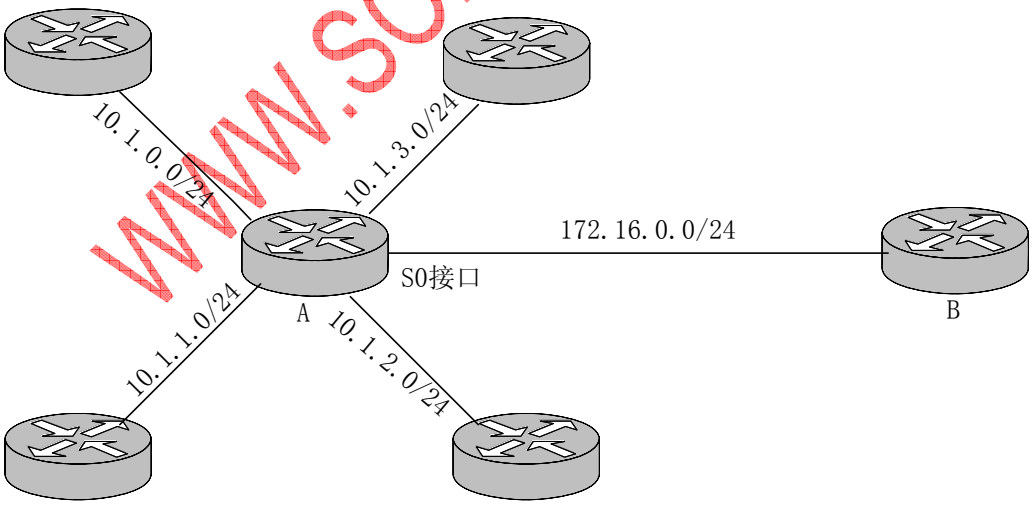
172.16.2.10000000到172.16.2.11111111两个子网.

被分的那个网段,已经不再是一个网段了，而是多个网段了，所以不能再分配给用户,用户只能用由它分出来的更小的子网网段.

2 SUMMARY汇总:减少路由表中条目数目的技术,用单条路由条目替代多条条目(通过寻找最大共有位来实现的)

R 10.1.0.0/24 即00001010.00000001.00000000.00000000
R 10.1.1.0/24 即00001010.00000001.00000001.00000000
R 10.1.2.0/24 即00001010.00000001.00000010.00000000
R 10.1.3.0/24 即00001010.00000001.00000011.00000000
前22为相同,手动汇总到R 10.1.0.0/22
对一个地址来说,最多只能汇总到主类边界,上述的最多到10.0.0.0/8

自动汇总:直接汇总到主类边界,是不精确的，汇总过大，不好（例：RIP的不连续主网分配）
手动汇总:是精确的（classless支持手动汇总，classfull不行。）



RIP V2、EIGRP在接口汇总是减小与该接口相连的邻居路由器的路由表大小

```
A:router rip
    Version 2
    No auto-summary
    Network 10.0.0.0
  Int s0
    IP rip summary-address 10.1.0.0   255.255.252.0    (手动汇总到/22,在A的接口S0上做汇总，能减小B的路由表大小。)
```

3 CIDR互联域路由:突破主类边界,进一步汇总.(主要用在互联网的C类地址,在局域网不用,它已**没有主**

类地址的概念的了)

202.103.24.0/24　202.103.25.0/24　202.103.26.0/24……汇总到202.0.0.0/8,这个已经那一类都不是了,所以是无类别的了.

4 查询路由表方式:classfull classless方式(以前说的classfull classless是路由协议,与此不同)

classfull:先看是否有主类匹配,如有看明细,无则丢弃.

Classless:不看是否有主类匹配,执行最大匹配查询(只有它支持CIDR查表)

10.0.0.0/8

10.1.1.0/24

10.1.0.0/16

10.0.0.0/8

10.1.1.1与10.1.1.0/24有24位相同,是最多的,所以匹配

命令:IP classless (IOS10.3版本后都默认该方式.)

IP classfull

5 defult route缺省路由

命令:IP route 0.0.0.0　0.0.0.0　x.x.x.x(下一跳,可以是电信局的路由器的接口)

在路由表中将出现这样的条目:　* s 0.0.0.0/0 x.x.x.x(下一跳)

只有RIP支持缺省路由的广播、传播。eigrp、ospf都不支持，要重分布。它不广播非本协议network生成的条目，包括静态的缺省路由以及是本协议的而未经network指定的。

6 层次化的地址分配：分配地址时，要能汇总起来。

EIGRP:是个杂合的,classless,IGP,的cisco私有的动态路由协议.

1 路由信息只能在邻居间交换(邻居是由一个子网直连的两个路由器)

邻居表(EIGRP neighbor table) 注意:它与路由表不同.

查看邻居表指令:show ip EIGRP neighbor

2 HELLO包:作用 1.探索邻居 2.跟踪邻居的存在.

3 HELLO包的间隔时间:1.当该路由器接口速率>T1=1.544Mb/s时,间隔为5秒.

2. 当该路由器接口速率<T1=1.544Mb/s时,间隔为60秒

4 holdtime:抑制时间

HELLO包中的一个递减的时间字段,在邻居表中为每个邻居单独维护的时间,当递减到0时,从邻居表和路由表相关条目都被删除.

Rip的holdtime是当收到条目的更新跳数大于原有的，则在180秒类保持该条目不变。

Holdtime=3倍的HELLO周期时间 (15S或180S)

5 邻接:交换路由信息(update)的过程.(交换的是真正的路由网段信息)

EIGRP第一次交换的是完整路由信息,以后就是只交换有改变的部分.(所以减少了带宽的占用)

6 拓扑结构表(topology table)

7 EIGRP是属于classless的动态路由协议.所以它支持VLSM(可变长掩码)和不连续子网划分.

8 EIGRP组播:组播地址 224.0.0.10

9 EIGRP的最大跳数为255跳,缺省值是100跳.

10 EIGRP具有第四层的功能,传输层的功能:传输的可靠保证.

1 发送对象是邻居.

2 发送的包是有序号标记的

3 接受方以ACK应答来保证接收到并通知发送者.

组播如没ACK回应,包会单播重发,重发次数最高是16,如还无ACK回应,则会从邻居表中将其删除,并将路由表中相关条目删除.

其窗口大小固定为:1，没有滑动窗口的概念，即没有流控的概念。

11 EIGRP共有5包:Hello　ACK　Update　query查询 reply答复(与路由无关的包就不考虑传输的可靠性)

12 metric值是由5个值计算得来的,包含了带宽信息.

Metric的计算:沿路的最低带宽+沿路的总延迟

13 dual算法:计算拓扑结构表得出路由表(它的算法可避免路由环路,而且可以加快收敛速度)

dual的计算:FD:自己到目的地最小metric值

RD:邻居到达目的地的metric值(邻居告诉你的)

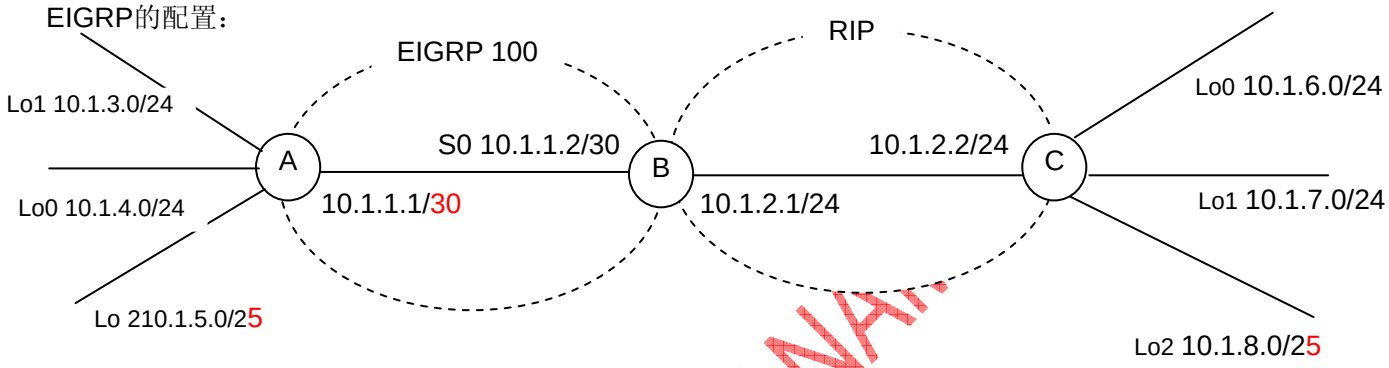
FC:RD(邻居)＜FD(自己最小)条件被满足否，是指目的地到自己的最后一跳是这个邻居，所以到目的地的条目下一跳就是这个路由器，也可以说邻居到目的地比自己近，从而避免了routing loop.

S:自己到达目的地的最佳路径.(记录到路由表中的)

FS:满足条件FC路径(备份路径:放入拓扑结构图中,最佳路径S失效时，经计算后进入路由表替代S)

环路可能出现的条件:邻居到一个地点的metric值(AD)比它先前告诉你的值变大了

EIGRP的配置：



A:router eigrp 100

No auto-summary

Eigrp log-neighbor-change (记录邻居状况变化过程，用show logging查看)

Network 10.1.0.0 0.0.255.255 (它含盖了A的4个子网，0表示必须匹配，1表示任意，与掩码是反码的关系)

Passive Lo0 不向用户网段Lo0发送路由信息

Passive Lo1 不向用户网段Lo1发送路由信息

Passive Lo2 不向用户网段Lo2发送路由信息

B：router eigrp 100

Eigrp log-neighbor-change

Network 10.1.1.2 0.0.0.0

Router rip

Network 10.0.0.0

Version 2 只接收、发送rip v2的路由更新包

No auto 关闭自动汇总

Passive s0 不向网段s0发送rip类的路由信息

C:router rip

{ Version 2 } 这两行是成对出现的，因为rip v1的自动汇总默认是开的。
{ No auto }

Network 10.0.0.0

Passive Lo0

Passive Lo1

Passive Lo2

重分布：在一个接口学到的一种路由协议的条目，从处于不同协议的接口发出。

B：router rip

redistribute eigrp metric 2 把eigrp100的条目放到rip中去。并设置初始跳数2

router eigrp 100

Redistribute rip metric 10000 100 255 1 1500 把rip的条目放到eigrp100中去。并设置初始metric的参数：带宽10000、延时100、可靠性255（100%）、负载为1（无负载）、MTU为1500。

（以太网的帧最大为1518byte,最小为64byte,IP包最大为65535byte,因为IP包大于帧，所以要把IP包分别封装为若干个帧，这个帧的最大字节数就是MTU表示的字节数。）

如果上图的A路由器上再加4个loopback接口，30.1.0.0/24、30.1.1.0/24、30.1.2.0/24、30.1.3.0/24，可以用手动汇总来减少路由条目：

A:int s0
ip eigrp summary-address 100 30.1.0.0 0.0.3.255
ip eigrp summary-address 100 30.1.0.0 255.255.252.0 (待定，不知道用那个)
实际就指是30.1.0.0/22的网段，划分它得到的子网包含了这4个子网。

自动汇总，是不支持不连续的主网IP地址分配的，所以A上要关闭自动汇总。

不汇总，路由条目多，占用CPU、内存、带宽等资源

试验用到的其它命令：

Ip host B 10.1.1.2 （绑定ip与B的telnet,在输入过这个命令的机器上到B去，只用输入B就可以了。）

Show ip eigrp neighbor （查邻居表）

No ip domain-lookup （关闭域名查询。）

Clear ip router * （马上收敛）

OSPF :所有厂家都支持，classless动态路由协议，IGP类的，会生成全网拓扑、完整的链接状态的协议。

1. neighbor table,看该表：show ip ospf neighbor
2. hello包，间隔周期10秒。作用：探索邻居，维护邻居的联系。
   Hello包的作用：发现邻居、跟踪邻居、选DR和BDR.
3. holdtime,邻居表中的为每个邻居设的。
4. 邻接，交换有关路由信息的过程，不是交换路由条目。
   邻接的条件：
   - 两者的hello time、holdtime要一致。
   - Area-ID要相同
   - Stub area标记要相同
   - Password要相同

Router-ID:是 OSPF 邻居表的字段,表示邻居的标记，可以是与其他 router－ID 都不同的任意 32 位的正整数。.是在 OSPF 域中唯一标记某个运行 OSPF 的路由器。因为一个 IP 是唯一的，所以惯用 IP 地址做 router-id

- Loopback 接口的 IP 地址是被优先选中成为 OSPF 的 router-ID.比以太接口优先级别高.
- 在没有设定 loopback 接口的情况下,在以太接口 IP 地址中,数值最大的被选中,来做为 router-ID.
- 一般人为的用 loopback 来固定表示一个 OSPF 路由器的 router-ID.(因为如果当做为 router-ID 的以太接口 down 了,在路由器重启后或 clear ip ospf process 后，该路由器的 router-ID 会重新选择.)
- 可以手动设 router-ID,如:router-id 192.168.1.1

OSPF interface network type(OSPF 接口网络类型):决定着邻居接口之间是否建立邻接,以及如何邻接.与 eigrp 不同,eigrp 是只要是邻居就邻接。

原始类型有 3 类：

一. Point-to-point 类型: 2 层封装协议为 HDLC、PPP 格式帧的接口。
邻接**肯定**建立。

二. BMA（广播多访问网络）类型：如以太网、令牌网。

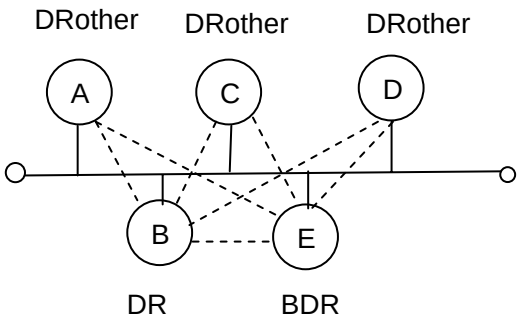
建立邻接的过程：（由 hello 实施的）

1. 选一个路由器为 DR，大家都和 DR 连接。（减少了不必要的网络连接）
2. 另外选一个 BDR 为 DR 的备份，当 DR 失效的时候，BDR 启用为 DR.（大家也和 BDR 邻接,且 DR 与 BDR 也连接）
3. 其他的全为 Drother.

最后的状况是 Drother 都和 DR、BDR 连接，BDR 与 DR 连接。

BMA 中接口的 Priority(优先级)：范围是 0～255，缺省值为 1。大的选中。如 priority 相同，router-id 大的选中。0 是永远不会被选为 DR、BDR。当新加入的 router 的接口 priority 更高，其中 DR、BDR 会不变的。除非 clear ip ospf process,所以打开 router 的电源时，按 priority 高低依次开

对于下图，eigrp 会邻接 N×（N-1）÷2 次。

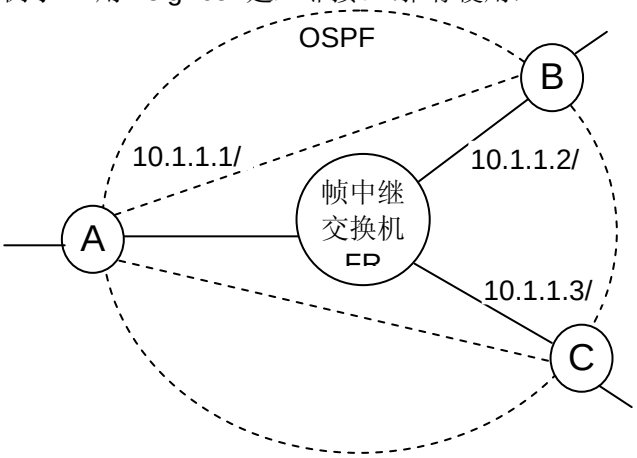


三. NBMA（无广播多访问网络）类型：2 层帧的封装协议为 X.25、帧中继、ATM 的接口。（邻接必须手动）

邻接方式有 4 种，推荐使用前两种，因为它们没有改变路由表。

1. 用 neighbor 建立邻接，但还要选 DR、BDR。（推荐使用）
2. 修改接口的网络类型，由 NBMA 改为 BMA。（推荐使用）
3. 修改接口的网络类型，由 NBMA 改为点到多点型 point-to-multipoint(不推荐)，然后会自动邻接。
4. 修改接口的网络类型，由 NBMA 改为点到点型 point-to-point（不推荐），然后会自动邻接。

例子 1 用 neighbor 建立邻接。（推荐使用）



启动 ospf

A:router ospf 1　　(这里的 1 是针对 A 这单个路由器的进程号)
　Network 10.1.1.1　0.0.0.0　area　0
B:router ospf 1
　Network 10.1.1.2　0.0.0.0　area　0
C:router ospf 1
　Network 10.1.1.3　0.0.0.0　area　0

建立邻接，及设定 DR

A: Neighbor 10.1.1.2　　(A 与 B 邻接，B 上不用再用 neighbor 指令了)
　Neighbor 10.1.1.3　　（A 与 C 邻接，C 上不用再用 neighbor 指令了）
　Int s0
　　Ip ospf priority 255　　(A 的 s0 接口被人为的设为 DR)
B:int s0
　　Ip ospf priority 0　　(B 的 s0 接口永远不会成为 DR 或 BDR)
C:int s0
　　Ip ospf priority 0　　(C 的 s0 接口永远不会成为 DR 或 BDR)

例子 2：修改接口的网络类型，由 NBMA 改为 BMA。（推荐使用）

A:int s0
　Ip ospf network broadcast　　(将 A 的 s0 接口改为 BMA 广播多访问型)
B:int s0
　Ip ospf network broadcast　　(将 B 的 s0 接口改为 BMA 广播多访问型)
C:int s0
　Ip ospf network broadcast　　(将 C 的 s0 接口改为 BMA 广播多访问型)

(用 show ip ospf interface s0 查看 s0 接口的网络类型)

A:int s0
　Ip ospf priority 255 (将 A 定为 DR)
B:int s0
　Ip ospf priority 0　(B 永远都不会成为 DR 或 BDR)
C:int s0
　Ip ospf priority 0　(C 永远都不会成为 DR 或 BDR)

例子 3：修改接口的网络类型，由 NBMA 改为点到多点型 point-to-multipoint(不推荐)。

A:int s0

Ip sopf network point-to-multipoint (将 A 的 s0 接口改为 point-to-multipoint 型)

B:int s0

Ip sopf network point-to-multipoint (将 B 的 s0 接口改为 point-to-multipoint 型)

C:int s0

Ip sopf network point-to-multipoint (将 C 的 s0 接口改为 point-to-multipoint 型)

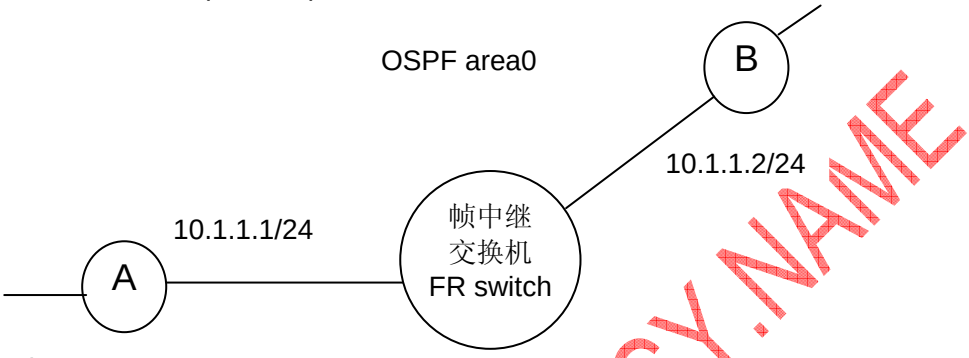
A 路由器条目相应会多出 2 个：10.1.1.2/32 和 10.1.1.3/32 （主机路由）

B 路由器条目相应会多出 2 个：10.1.1.1/32 和 10.1.1.3/32 （主机路由）

C 路由器条目相应会多出 2 个：10.1.1.1/32 和 10.1.1.2/32 （主机路由）

例子 4：修改接口的网络类型，由 NBMA 改为点到点型 point-to-point（不推荐）

上图不能改为 point-to-point（因为有 2 个以上的接口参与），下图可以



A:int s0

Ip sopf network point-to-point

B:int s0

Ip sopf network point-to-point

Ospf 其他的网络类型接口：

Loopback 接口类型：它的路由条目掩码是 32 位的，除非改该接口为 point-to-point 后，可使其掩码变为 24 位的了，意义不大。

概念：

LSA：链接状态的 IP 广播，它描述了全网链接状态的数据结构。OSPF 交换的是 LSA,不是路由条目。

运行 OSPF 的路由器第一次交换的是完整 LSA 信息，以后交换变化的部分。（IGP 类的动态路由协议最终要形成全网路由表，这里是形成全网拓扑，再用 SPF 算法算出全网路由表。）

LSA database:每个路由器都有一个 LSA database,大家的 LSA database 收敛后就全部一样了，同步就实现了。

SPF 算法：最短路径优先算法，每个路由表都用 LSA database 以自己为根，算出到每个网段的最短路径的树，形成全网路由表。

网络稳定的时候，网中只有 HELLO 包流动，无 LSA 传递。EIGRP 也是的。

邻居状态机制（邻接建立的过程）：注意，这里的状态都是说的是邻居的状态。

1. 我收不到邻居发送的 hello 更新包，就认为邻居为 down 状态。
2. 我收到邻居发的 hello 包，就认为邻居处于初始化状态。
3. 在这第一次发 LSA 时，相互把全部邻居表 neighbor table 发给邻居，我从邻居发来的 neighbor table 中提出我没有的 neighbor table 条目，填充到我的 neighbor table 中去。
4. 我能收到邻居 hello 包，并在邻居的 hello 包的 neighbor table 中，发现有自己的 router-ID,则该邻居达到 two-way 状态。
5. 在该邻居 two-way 状态下，判断其接口网络类型：piont-to-point BMA NBMA 是属于哪一类。

6. 如为 BMA，在判断 DR,BDR,Drother 后，邻居处于 exstart 状态。所以 Drother 保持在 two-way 状态下。
7. 在 exstart 状态后，邻居和我的接口选定主从关系（只是为了方便控制两者的交换信息而已），与选 DR 不是一回事。（router-id 数值大的做为主。）
8. 进入 loading 状态：
   这其中 DD:LSA 的摘要信息 ； LSR：自己不知的邻居而有的 LSA 摘要的集合 ；LSU：对应 LSR 的那些真正的 LSA 条目。
   a) 主，发 DD,只发一次（DD 是 LSA database 的 LSA 摘要字段的全部集合，即头信息的全部集合），邻居达到 exchange 状态。
   b) 从，发 LSR（LSR 请求发送的集合，是要求主发送，**“从”没有而“主”有**的 LSA 条目的摘要集合）（R：request）
   c) 主，发 LSU(LSU 对应“从”LSR 请求的那些 LSA database 的完整条目信息)(U：update),
   d) 再反之，从发 DD，主发 LSR，从发 LSU。
9. full state(状态):邻接建立完成。这时，所有路由器的 LSA database 都相同了。

建立邻接后，就只有 hello 包在邻居间交换了。

OSPF 划分为若干个区（area），一个 area 内的所有路由器维持着一个单独的 LSA 表，在区内同步。

1. Area 的作用：使 LSA 减小，使路由器带宽、CPU、内存的占用减小。
2. 在划分了多个区时，area 0 必须存在，是称为 backbone 即主干区。（如果只有 1 个区，也可以命名为 area 1 或 area 2 等等），且其它 OSPF 的区必须和 area 0 区连接。（或者通过 virtral-link 技术做成的传输区和 area 0 链接。）这样是为了防止路由环路，区间是距离矢量，所以有可能出现环。区边界路由器 ABR 把一个区的 LSA 发到区 0 时，有个源起者说明是自己，当穿过 2 个区时，源起者就变了。所以区必须连区 0。
3. virtral-link 技术：把一个非 area 0 的区变为传输区，该区两边分别和 area 0 及另一个非 0 区连接，充当两个区的传输区（把该区做成了属于 area 0 的一条网段一样）。

OSPF 路由器的分类：

1. Internal（内部路由器 IAR）：这个路由器的接口处于同一个区内，所以该路由器只需维护此区的 LSA database.
2. ABR（区边界路由器）：至少有接口从属不同的区（所以必须有和 area 0 相连，或者是传输区？？），所以要维护它参与的每一个区的 LSA database.
3. ASBR 自治系统边界路由器：（这里是书上定的名字，不准确，应该是 OSPF 域边界路由器更准确，这里并不是指运行不同动态路由协议的域的集合的自治系统。）有接口处于 OSPF,还有接口处于运行其他动态路由协议的域中，其间要进行动态路由重分布。
4. backbone:至少有一个接口处于 area 0 中。

LSA 的分类：

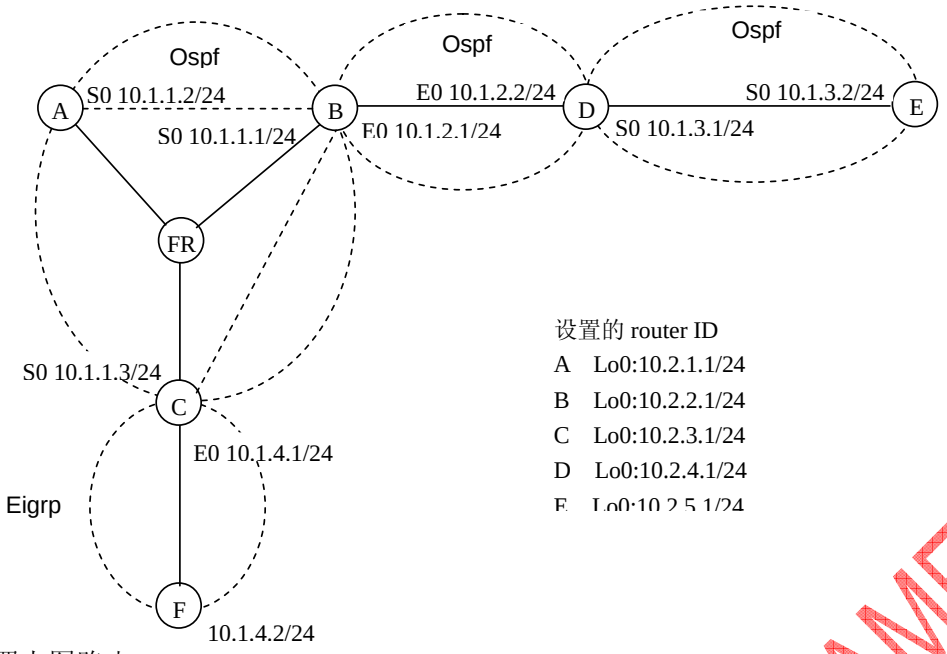
1. LSA 1(router LSA):每台路由器都会为它的每个接口产生一个 LSA,它必然从属于一个区，它不能穿越区，是最基础的 LSA，为生成区内全网拓扑。.
2. LSA 2（network LSA）:DR 产生的，用来描述它所连接的 BMA 网络，它只在区内蔓延，不能穿越区，也是基础的 LSA.
3. LSA 3(network summary LSA)：ABR 将一个区的 LSA 1、LSA 2 发到另一个区，在整个 OSPF 域内传递（除了产生该 LSA 的那个区），变为 LSA 3。
   区汇总 area summary（手动的）:可有效的减少该区 LSA 3 的数目。区汇总 LSA 不再含有连接信息，不能推出连通图的，只能从这种 LSA 上知道包含了哪些网段。
   OSPF 区间的路由是距离矢量路由，cost 累加。
   OSPF 区内的路由是链接状态路由，spf 算法。
4. LSA 4（ASBR summary LSA）:ABR 产生的，告诉如何到达 ASBR 的 LSA 条目（所以生成的路由条目是 32 位掩码的主机路由）。在整个 OSPF 域内传递（除了产生该 LSA 的那个区）。

5. LSA 5（external LSA）外部 LSA:ASBR 产生的，告诉如何联系外部不同协议域网段的路由，为每个外部路由产生一个 LSA5. 在整个 OSPF 域内传递。
   外部路由汇总:可有效的减少 LSA 5 的数目
6. LSA 7(NSSA 区):

routing loop 出现的可能分析：
OSPF 在区内没有，因为 SPF 算法。
OSPF 在区间没有，因为须连区 0 的设计。
在 OSPF 域外，LSA5 产生的路由条目，有可能有路由环路，因为 OSPF 没有措施防范。

不同 LSA 在路由表的表示：
1. O :LSA 1 , LSA 2 生成的。
2. O IA (Inter Area 区间路由):LSA 3 ,LSA 4 生成的
3. O E (external 外部路由)：LSA 5 生成的
   a) O E1:外部到 ASBR 的代价 cost（初始 cost） + 自己到 ASBR 的 cost
   b) O E2: 外部到 ASBR 的代价 cost（初始 cost）O E 条目的缺省值
4. O N （NSSA）：LSA7 生成的。

区的分类：
1. backbone 区,即主干区
2. 标准区：内部允许有 LSA 1~LSA 5 的 5 种 LSA
3. stub 区：内部有 LSA 1~LSA 4 的 4 种 LSA，没有 LSA 5,即无外部 O E 路由。由 ABR 自动地发出缺省路由，让区内的路由器来连接外部的路由器。
   a) backbone 区、传输区不能做成 stub 区
   b) 只有一个 ABR 的且没连外部网络的边缘区，适合做成 stub 区。
   c) stub 区不能连接外表路由域（因为它不传递 LSA5）
4. total stub 区（cisco 私有的）：内部有 LSA 1 ,LSA 2 ,没有 LSA 3~LSA 5.由 ABR 自动广播，发出缺省路由。
   a) backbone 区、传输区不能做成 total stub 区
   b) 只有一个 ABR 的且没连外部网络的边缘区，适合做成 total stub 区。
   c) total stub 区不能连接外表路由域（因为它不传递 LSA5）
5. NSSA 区：ABR 上会阻止 LSA5 从区 0 注入本区，ASBR 进外部 LSA 用 LSA7，ABR 会把 LSA7 翻译成 LSA5.(我自己忘了这句话的本意了)



设置的 router ID

A Lo0:10.2.1.1/24
B Lo0:10.2.2.1/24
C Lo0:10.2.3.1/24
D Lo0:10.2.4.1/24
E Lo0:10.2.5.1/24

配置上图路由：

B:router ospf 1 （进程 1，针对路由器本身来说的，邻居间进程号可以不一样。是在同一个路由器上区分不同 OSPF 的域）

Router-id 10.2.2.1 (手动设 router-id,可以不要，而去设 loopback 0)
Network 10.1.1.1 0.0.0.0 area 0
Network 10.1.2.1 0.0.0.0 area 1
Network 10.2.2.1 0.0.0.0 area 1 (B 的 Lo0 被放入非 0 区，是为了减少主干区的路由运算)
Passive Lo0 (不向用户网段发 hello 包)

A:router ospf 1
Router-id 10.2.1.1 (手动设 router-id，可以不要，而去设 loopback 0)
Network 10.1.1.2 0.0.0.0 area 0
Network 10.2.1.1 0.0.0.0 area 0 (A 的 Lo0)
Passive Lo0 (不向用户网段发 hello 包)

C:router ospf 1
Router-id 10.2.3.1 (手动设 router-id，可以不要，而去设 loopback 0)
Network 10.1.1.3 0.0.0.0 area 0
Network 10.2.3.1 0.0.0.0 area 0 (C 的 Lo0)
Passive Lo0 (不向用户网段发 hello 包)

B:router ospf 1
neighbor 10.1.1.2
neighbor 10.1.1.3
int s0
ip ospf priority 255 (提高 B 的 s0 的优先级，使 B 的 s0 成为 area 1 的 DR)

A 和 C:
int s0
ip ospf priority 0 (当 B 的 s0 成 DR 后，A、C 的 s0 会试图生成 BDR，用此命令使他们任何一个都不可能成为 BDR)

D:router ospf 1
Network 10.1.2.2 0.0.0.0 area 1
Network 10.1.3.1 0.0.0.0 area 2
Network 10.2.4.1 0.0.0.0 area 0 (D 的 Lo0)
Passive Lo0 (不向用户网段发 hello 包)

B：int e0

Ip ospf priority 255 (设 B 的 e0 为 area 1 的 DR)

E:router ospf 1

Network 10.1.3.2 0.0.0.0 area 2

Network 10.2.5.1 0.0.0.0 area 2 (A 的 Lo0)

Passive Lo0 (不向用户网段发 hello 包)

B:router ospf 1
Area 1 virtual-link 10.2.4.1 (10.2.4.1 是 D 的 router ID,成对的发，D 上也发)
D:router ospf 1
Area 1 virtual-link 10.2.2.1 (10.2.2.1 是 B 的 router ID)
C:router eigrp 100
Network 10.1.4.1 0.0.0.0
No auto-summary （关闭原来缺省的自动路由汇总，OSPF 是没有自动汇总的）
Eigrp log-neighbor-change (记录下 C 的 eigrp 的邻居变化过程，方便找错。)
F:router eigrp 100
Network 10.1.4.2 0.0.0.0
No auto
Eigrp log-neighbor-change (记录下 F 的 eigrp 的邻居变化过程，方便找错。)
C:router eigrp 100
Redistribute ospf 1 metric 10000 100 255 1 1500 (把 ospf 重分布到 eigrp 100 去)
router sopf 1
redistribute eigrp 100 subnet metric 30 (把 eigrp 100 重分布到 ospf 去,)
(只有重分布到 ospf 时，metric 有缺省值 20 存在（只有 OSPF 才有缺省值），这里我们改成 30)
（subnet,如果不加它，重分布过来的只有主类网段，不含其中的子网）
redistribute eigrp 100 subnet metric－type 1
(OE 条目缺省值为 OE2,即不加自己到 ASBR 的 cost，用 metric－type 1 就改外部路由为 OE1 类型了)

假设 D 上又多了 4 个网段：
Lo10:172.16.0.0/24
Lo11:172.16.1.0/24
Lo12:172.16.2.0/24
Lo13:172.16.3.0/24
D:router ospf 1
Network 172.16.0.0 0.0.255.255 area 1
或 redistribute connected [metric 20] subnet （也可以用重分布把这些网段引入 D 的 OSPF 中,connected 是直连的意思）

B:router ospf 1
Area 1 range 172.16.0.0 255.255.252.0
还是 summary-address 172.16.0.0 255.255.252.0 ???
(在 B 上手动汇总（找这些网段的最大共有位），减小 A、C 的路由表，而不是减小 B 的路由表。)
（注意：eigrp 是在接口上作汇总，ospf 只能在 ABR、ASBR 上做区内汇总）

把区 2 做成 stub 区。
D 和 E:router ospf 1
Area 2 stub （要在 area 2 中的每个路由器上都发出此命令）

或，把区 2 做成 total stub 区。
D:router ospf 1
Area 2 stub no-summary (只要在 area 2 中的唯一的 ABR 路由器上发出此命令)

查看 virtual-link:的命令：

show ip ospf virtual

(显示了 adjacency state full (hello suppressed)才说明 virtual 成功)

查看 eigrp 邻居状况变化过程记录的命令：

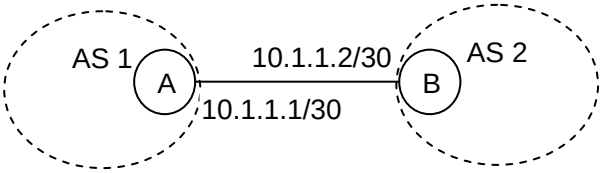
show logging

相当于 reset: clear ip ospf process

BGP:自治系统 AS 之间的动态路由协议。一个 AS 可以一个省级、一个国家的网，必须控制 AS 间的路由交换。所以于 IGP 生成全网拓扑的方式不一样。

BGP 涉及概念：

1. 自治系统号是由 NIC 组织分配的，1～64511 分配给了互联网，64512～65535 为私有。
2. Neighbor:
   a) BGP 的邻居可以跨越路由器，即 BGP 邻居间可以隔着其它的路由器。只有 IP 地址能 PING 通，就可以手动的建立一个邻居。
   b) BGP 的邻居必须手动的建立，不是 hello 包探测得来的。
3. 路由信息的交换：BGP 的邻居建立以后，不自动地交换任何信息。与 IGP 类的完全不同，IGP 类的会自动交换信息直至生成全网拓扑。
4. BGP-table:BGP 的所有操作都是围绕 BGP－table 来做的。最初始时，BGP－table 是空的。



5. 在 BGP-table 中手动添加条目，有 3 种办法，唯一被推荐的是用命令 network

```
A:router bgp 1
    Neighbor 10.1.1.2 remote-as2
B:router bgp 2
    Neighbor 10.1.1.1 remote-as1
```

手动地建立邻居

可以用 show ip bgp neighbor 来查看邻居是否建立起来了

用 show ip bgp 查 BGP 表中的内容。

BGP 只传递主类汇总，或 CIDR(汇总突破主类边界的：超网)

A:network 10.0.0.0 mask 255.0.0.0 这里假设了 AS1 中全是 10. 0. 0. 0/8 内的子网。

BGP 协议中，network 只是为了在 BGP table 中添加条目，没其他的作用。这于 IGP 类的是完全不一样的（network 在 IGP 的作用是：1 参与此协议的接口有那些。2 该接口的 IP 地址的网络地址也随更新包发出去了）

对于 A 的 BGP table 来说，如果 A 的路由表中没有 10.0.0.0/8 的条目，则 A 的 BGP table 中则不能出现 10.0.0.0/8 的条目。所以要在 A 的路由表中静态生成路由条目 10.0.0.0/8,A 的 BGP table 中才会有 10.0.0.0/8 的条目出现

A:ip router 10.0.0.0 255.0.0.0 null 0

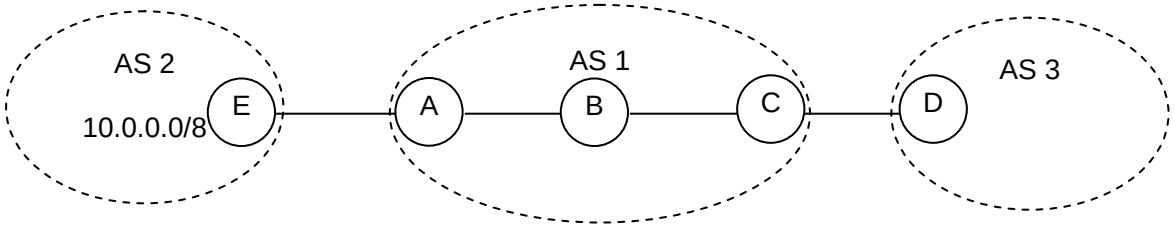
Null 0 是垃圾桶接口，所有发往 null 0 的包都会被丢弃。当查询时，与前面的条目都不匹配时，就只能和 10.0.0.0/8 匹配了（这里前提是 AS 1 内全是 10.0.0.0/8 内的子网），就被丢弃了。这是为了防止路由环路出现的可能。

因为邻接已经建立了，这时，A 会把这个 BGP table 中的条目广播给 B。

BGP 的广播与 IGP 类的广播、组播不同。BGP 广播是调用 TCP 会话（端口为：179），是基于 3 层的 IP 地址和会话层的端口的，是点到点的会话。是可靠的传输。对于 IGP 的 OSPF、EIGRP，它们自身有可靠传输的组播发送机制（有个字段用于发送的可靠）。

TCP 是面向连接的服务

UDP 是面向无连接的服务。



注意：B 没有参与 BGP

BGP 邻居的两种关系：

1. EBGP:如果 BGP 邻居从属于不同 AS，他们的关系则为 EBGP。如这里的 E、A.
2. IBGP:如果 BGP 邻居同属于一个 AS，他们的关系则为 IBGP。如 A 和 C

（有了 A 和 C 这两个 IBGP,AS 2 和 AS 3 就连接起来了，就象 AS 1 是一个传输 AS 一样。）如果没有 IBGP，AS 2 到 AS 3，就要在 A 上作 BGP 到 AS 1 内的 IGP 重分布，A 的 IGP 路由表就有了一个 B 的条目，到 C 时，再从 C 上作 IGP 到 BGP 的重分布。这是非常不好的，IGP 到 BGP 的重分布是不安全的，所以是不被允许的。

BGP 的同步：一个运行 BGP 的路由器，不会把从一个 IBGP 路由器学到的 BGP 条目网段传递给一个 EBGP. 当这个运行 BGP 的路由器的路由表没有该网段存在的时候。在该网段在路由表中出现了，就传递给它的 EBGP。

结合上图来看：一个运行 BGP 的路由器(C)，不会把从一个 IBGP 路由器(A)学到的 BGP 条目网段（10.0.0.0/8）传递给一个 EBGP(D).当(C)这个运行 BGP 的路由器的路由表没有该网段存在的时候。在该网段（10.0.0.0/8）在（C 的）路由表中出现了，（C）就传递给它的 EBGP（D）。

就是说，BGP 路由器 C 的 BGP table 和路由表都要有 10.0.0.0/8（这个从它的 IBGP 邻居 A 学到的网段），C 才会把 10.0.0.0/8 这个网段 BGP 广播给它的 EBGP 邻居 D。不满足这个条件时，C 是不会把该网段 GBP 广播给 D 的。

因为，AS3 到 AS2 时，不同步的话，包就会在 B 处丢失了。（路由黑洞，在 IGP 表中走不通）

因为，B 是没有参与 BGP，所以 B 的 IGP 路由表中没有自治系统外的 AS 2 的网段 10.0.0.0/8 的存在。

所以，要在 A 上做 BGP 到 IGP 的重分布，C 才达到同步的条件，这样 B 才知道 AS 2 中 10.0.0.0/8 这个路由的路径。

（假设 A 的 IGP 用的是 OSPF）

```
A:router ospf 1
    Redistribute BGP2 [metric 20]
```

BGP 的同步是一个开关，缺省是开同步检测的，可以关同步检测：no sync

达到同步要求的办法：

1. 如上做重分布
2. 静态（在 C 上做静态路由以达到同步条件，但是还是会出现路由黑洞，所以同步检测的目的没达到。除非再在 B 上做 B 到 A 静态路由）

BGP总结：(BGP邻居间，第一次发完整条目，以后发变的部分)

被选中的含义：1、该BGP路由条目会进入IGP路由表。2、该BGP路由条目会发给被允许的BGP邻居。

BGP无并行路由，只选择1条最佳路径，且该路由条目会被选中。

我**从**哪学的	我会**发**给谁
从EBGP邻居学到的路由	被选中（进IGP路由表），并发给它的EBGP、IBGP邻居
从IBGP邻居学到的路由	不会发给另一个IBGP的（水平分割）
	能否发给EBGP，取决于同步检测条件

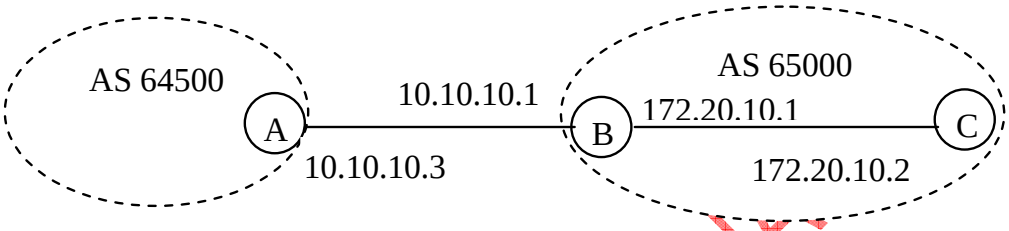
BGP其它涉及概念：

BGP属性按必须性分类：

必遵属性： AS-path、next-hop、origin
可选属性：Local preference attribut、MED、Weight、communication

BGP属性按传递性分类：
传递属性：AS-path、next-hop、origin、Local preference attribut、communication
非传递属性：MED、Weight

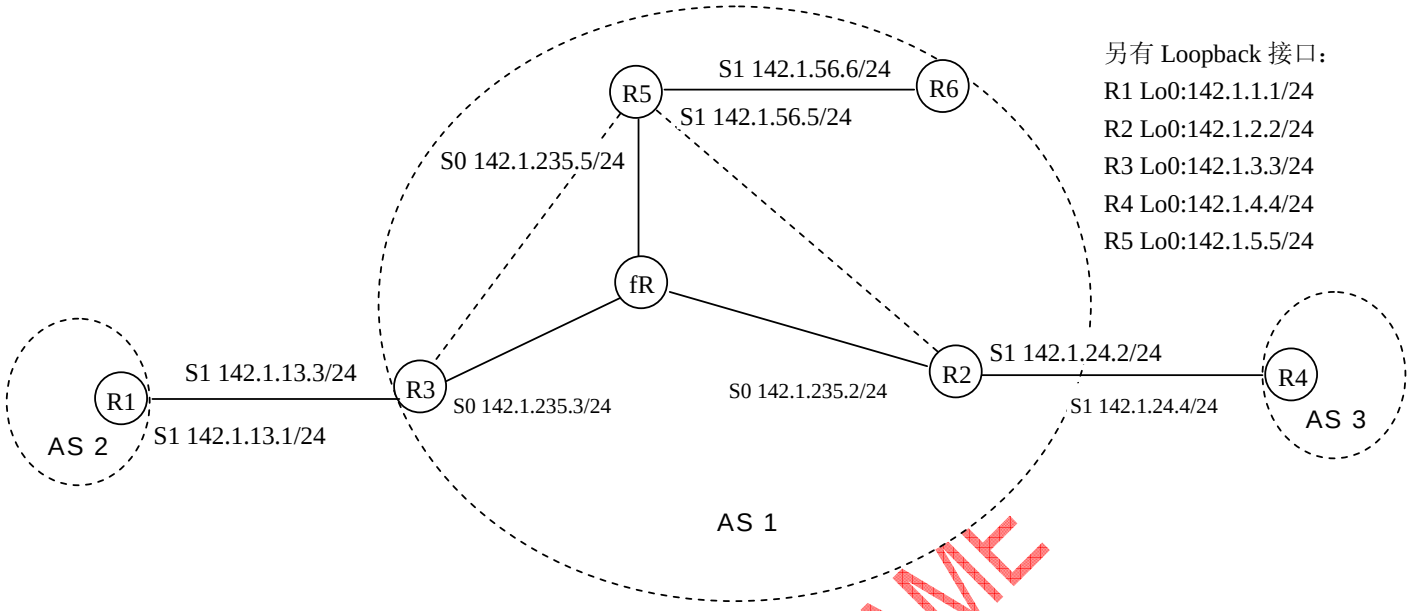
1. Metric,同一协议中，用来选择最佳路径的参数。
   如BGP:attribute , EIGRP：metric , OSPF:cast
2. AS-path ：从源到目的网段经过的AS的集合，短的优先。（包括经过的AS和目的AS，不包括源AS）
3. next-hop：BGP的下一跳next-hop在路由表中必须是可到达的（即BGP相关网段，在IGP的路由表中须有相关条目），否则该BGP条目永远不会被选中，所以不会进入路由表、不会发给BGP邻居。
   B:router bgp 65000
   Neighbor 172.20.10.2 remote-as 65000
   Neighbor 172.20.10.2 next-hop-self



4. BGP选取路径的参数：
   a) Local preference attribut(本地优选属性)：如AS有两个以上的出口，且出口不在一个路由器上时，选择出口的判断条件。
      i. 高的优选。
      ii. 该参数只在IBGP之间交换。
      iii. 在选择者上设定
   b) MED（或称metric）：如AS有两个以上的出口，但出口在一个路由器上时，选择出口的判断条件。
      i. 低的被优选。（这个参数类似代价cast）
      ii. 该参数只在EBGP之间交换
      iii. 在被选择者上设定。
   c) Weight宽度(cisco私有的)：功能和MED一样,为了改变MED的控制权在别人手里的问题。
      i. 高的优选
      ii. 该参数只在本地有效，不发给任何BGP邻居。
      iii. 在选择者上设定
   d) origin源起属性：
      i. i,BGP table条目是由network激活的，origin为I(类似IGP路由的方式)
      ii. e,BGP table条目是由EGP重分布过来的，origin为e(EGP的，AS间原来用EGP，BGP是EGP的升级)
      iii. e,BGP table条目是由直连和静态路由重分布过来的，origin为？（incomplete：不完善的）
      路径选择优先顺序：I ＞ e ＞ ?
   e) communication通讯：过滤路由作用。

BGP最佳路径选择比较顺序：
1. 比较下一跳是否可到达，不可到达的不能进入路由表。
2. weight高的优选，weight缺省是100
3. 本地优选属性高的优选，缺省值是0
4. AS－path短的优选。
5. origin(源起属性)：I ＞ e ＞ ?
6. MED低的优选（类似代价cast）
7. EBGP ＞ IBGP

8. 最佳路径，是指一条最佳的路径，所以最后只能选一条，如果有几条在上述步骤中分不出，就随机选一个。



R5:router bgp 1
　　Neighbor 142.1.235.2 remote-AS 1
　　Neighbor 142.1.235.3 remote-AS 1　　} 是IBGP
R2:router bgp 1
　　Neighbor 142.1.235.5 remote-AS 1　　(是IBGP)
　　Neighbor 142.1.24.4 remote-AS 3　　(是EBGP)
R3:router bgp 1
　　Neighbor 142.1.235.5 remote-AS 1　　(是IBGP)
　　Neighbor 142.1.13.1 remote-AS 2　　(是EBGP)
R4:router bgp 3
　　Neighbor 142.1.24.2 remote-AS 1　　(是EBGP)
R1:router bgp 2
　　Neighbor 142.1.13.3 remote-AS 1　　(是EBGP)
　　Network 200.200.200.0 mask 255.255.255.0　　（在R1的BGP表中源起200.200.200.0/24的网段）
　IP router 200.200.200.0　255.255.255.0　null 0　(在R1的路由表中生成该路由条目)
R3:router bgp 1
　　Neighbor 142.1.235.5　next-hop-self
　　Clear ip bgp　　(BGP只有在clear快速收敛后才会执行BGP命令，neighbor除外)
R5:router bgp 1
　　No sync　　(关掉同步)
也可在R3上重分布（BGP到IGP），以达到同步要求：
R3:router ospf 1
　　Redistribute BGP2 [metric 20]

BGP的水平分割：一个运行BGP的路由器，从一个IBGP邻居学到的路由，不会传递给另外一个IBGP邻居。
路由反射器技术：IBGP之间的技术
　　R2因为水平分割，不能从R5学到200.200.200.0的网段
　　Server: R5（象镜子一样）
　　Client: R2 R3（象光一样，被反射，传递给所有client,除了源IBGP）　　} 是IBGP之间的
　　命令格式：
　　R5:router bgp 1
　　　Neighbor 142.1.235.3　router-reflect-client　(含义：R5自己是server,R3被定为client)
　　　Neighbor 142.1.235.2　router-reflect-client　(含义：R5自己是server,R2被定为client)

不做全网状的拓扑，是为了不使BGP邻居太多而耗费昂贵的远程带宽。

R2:router BGP 1

Neighbor 142.1.235.5　weight 200　(R2所有从R5学到的路由的weight值由缺省值100变为200)

查看BGP邻居状态

show ip bgp neighbor　(邻居处于establish状态，那么邻居才是建立好了的)

查看BGP路由表

show ip bgp

router map类似访问列表，但它可以做操作修改（全局指令），主要用在BGP、重分布的过滤上。

格式：router map cisco　permit / deny　seq　(cisco是方案的名字，seq是序列号，缺省是5，指明查询顺序)

match ip address 访问列表号

……..

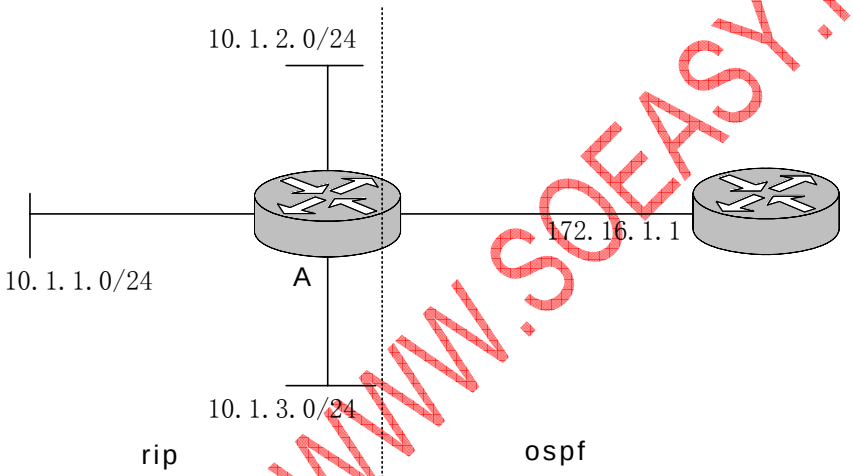
.

.

.

set…….　(一般改BGP属性)

- 这是一个节点，它可以有多个匹配语句，节点之间是“与”的关系。
- 和访问列表一样，最尾端会有一个隐含的语句拒绝所有的节点，所以节点都用permit.



重分布过滤网段10.1.1.0

A:router ospf 1

Redistribute rip router-map cisco　（cisco是router-map方案名）

Access-list 1 permit 10.1.2.0

Access-list 1 permit 10.1.3.0

Router-map cisco permit　（cisco是router-map方案名）

Match ip address 1　(1是访问列表号)

如果是 BGP过滤，可将前2行换成：

A:router BGP 1

neighbor 172.16.1.1　router-map cisco

DTP(动态trunk协议):在交换机之间自动生成trunk link.(cisco私有的)

tag:DTP自动帧标记，首选ISL标准,其次选dot 1q(缺省值)标准(用来区别VLAN号的标准)

它定义的trunk port接口模式：

1. on模式：如果相连的两个trunk port接口为on,它们之间会自动形成trunk link
2. descriptbe渴望模式：如果两个相连的trunk port为渴望模式，它们会自动生成trunk link;当一个on,一个descriptbe时，也可以自动生成trunk link.
3. AUTO模式：AUTO和AUTO不会自动生成，AUTO和ON、descriptbe渴望模式就可以
4. OFF模式：（端口关掉DTP，无DTP帧发出）两个trunk port,其中一个为off，就不可能成为trunk link.

Trunk port端口自动生成trunk link的情况：

	ON	Desirable（缺省）	Auto	Off
ON	生成	生成	生成	×
Desirable（缺省）	生成	生成	生成	×
Auto	生成	生成	×	×
Off	×	×	×	×

Trunk port缺省模式是：渴望模式Desirable.

VTP(Vlan trunk协议)：用来集中化管理VLAN配置的协议（cisco私有的）

它把交换机分为了3个角色模式：

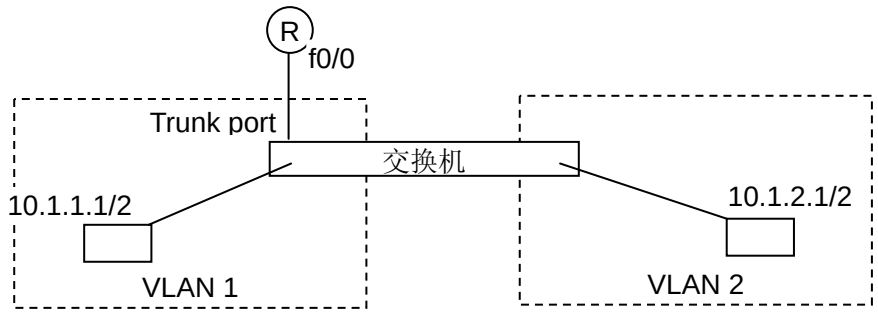
1. server模式（缺省）:能删除、添加、修改、生成VLAN，从trunk port发出广播VTP帧。一个域可以有多个server.
2. client模式:不能做对VLAN的任何操作。可以接受、转发VTP帧，同时查看自己的VLAN信息，同步更新自己。
3. 透明模式：也可以对VLAN做操作，但不广播出去。对于VTP帧，只接收和转发，但不同步更新自己，只是个通路作用。

VTP操作：

1. Server每改动一次VLAN，就会广播一个VTP广播帧（其中包括VLAN完全信息库），仅仅从trunk port广播出去。
2. VTP广播帧其中有个字段叫修订版本号（修订版本号是从0开始的，每当作出一个VLAN的改动，它就会加1的），用来判断VLAN的信息库是否更新过，是否是最新的。
3. VTP的域：一个管理的域，处于一个域（具有相同的域名）的交换机才会相互学习VLAN的信息，才会传递这个域的VTP帧。（一个域可以包括多个VLAN）

VLAN的互连方法：

1. 用路由器的两个接口各连接一个VLAN.
2. 单臂路由：



单臂路由

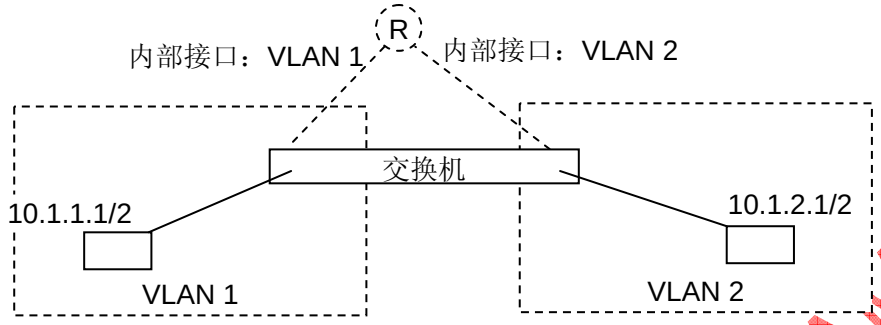
在路由器上的f0/0端口设：

```
int f0/0
  no ip address
  no shutdown
```

int f0/0.1
ip address 10.1.1.254 255.255.255.0
no shutdown
encapsulation isL 1 (封装 encapsulation 可简写为 encap，封装帧标记的标准用 isL 的方式，并将改口划为 VLAN1 的。这与把普通 access 接口划到 vlan 去不同)

int f0/0.2
ip address 10.1.2.254 255.255.255.0
no shutdown
encap isL 2 （封装帧标记的标准用 isL 的方式，并将改口划为 VLAN2 的)

这里用到了子接口，子接口：一个物理接口可以被划分为多个逻辑的子接口。原物理接口就不存在了。

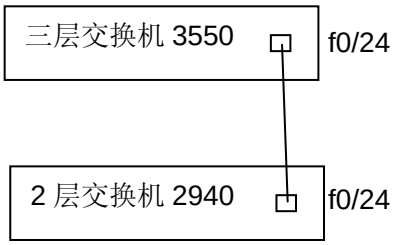


三层交换机互连 VLAN

3. 用三层交换机：(现在常用的方式)
三层交换机内置了一个路由器，
所以只需要进入相应VLAN的接口
设置它的地址就可以了
int VLAN 1 （命令执行后就自动生成了一个虚接口VLAN 1）
ip address 10.1.1.254 255.255.255.0
no shutdown
int VLAN 2
ip address 10.1.2.254 255.255.255.0
no shutdown
须有access接口属于vlan 1，否则虚接口vlan 1会down.

生成VLAN的步骤：

1. 建立trunk link。（一般是自动生成的）
如果要手动建立则如下：
3550:int f0/24
swi trunk encap dot1q
swi mod trunk (swi mod access是定端口为access)
2940:int f0/24
swi trunk encap dot1q
swi mod trunk



2. 定VTP域的名字，启动VTP功能
3550:VTP mode server (VTP模式默认是server模式，所以可以不写)
VTP domain cisco (这个VTP域的域名叫CISCO)
2940:VLAN database (这里与3550那个3层交换机不一样)
VTP client (VTP的模式为client模式，这里与3550那个3层交换机不一样，没有mode)
VTP domain cisco
3. 生成VLAN,在server上做
3550：VLAN 2
name 2
exit
VLAN 3

Name 3
Exit
等等…

删除vlan 2的指令: no vlan 2

如果2940做server,生成VLAN的命令有点不同

2940：VLAN 2　name 2

4. 为所有access接口划归VLAN(每个接口默认都是属于VLAN 1的。所以有必要时需手动划归VLAN)

如：

int f0/1
swi mod access
swi access VLAN 2

VTP是通过VLAN1（name是default）来传递系统信息实现的，所以VLAN1不能删、改（名）等，所以一般不把VLAN1作为用户VLAN，不把用户放入VLAN1

.

这里会用到的查看状态的命令：

查看端口模式的命令：

show trunk

要在接口上查看trunk link状态，用命令：

3550上：show int f0/24 trunk

2940上：show int f0/24 swi

看VTP广播帧的修订版本号

show vtp status

看VLAN的信息

show VLAN

注意：3层交换机3550路由功能默认是关闭的。打开路由功能命令：ip routing

cisco把交换机分为3类：

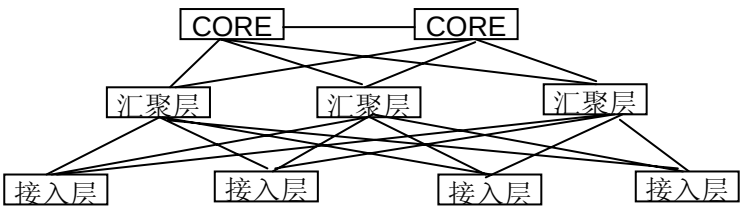
1. 核心层core：高速帧转发的角色。
   如cisco的6500：背板32G～256G，MAC表32K
   4500：背板可达32G，MAC表16K
   华为的8500：背板可达2.8T
2. 汇聚层（分布层）：是互连VLAN的角色，起路由的作用（如cisco的4500和3550）
   汇聚层到核心层的连接线路叫uplink
3. 接入层：是连电脑的角色（如cisco的3550、2940）

他们的互连称为uplink，用1G、10G的带宽。

在小的网络工程中，有时用汇聚层的设备充当核心core来实现VLAN互连路由，理论上的汇聚层就不存在了。

在重要的部门，为了网络正常工作，防止个别设备的失效带来的整个网络失败，用冗余技术。即用2个CORE核心层设备。做如下连接：

1. 核心层的互连
2. 汇聚层的每个都和2个CORE分别连
3. 接入层的每个都和每个汇聚层的连



广播风暴出现的原因：
1. 冗余必有环
2. 帧没有灭亡的时间限制字段
3. 透明桥接的广播功能

生成树协议（STP）：
作用：解决广播风暴
原理：把物理的环，变成逻辑的树。
其技术标准：IEEE802.1d
涉及概念：
1. BPDU（bridge protocal data unit）:是一个帧，为了计算STP而设计的，是在交换机间传递，是以组播的方式发送。
2. bridge－ID：BPDU中的字段，在STP域中唯一标识一台交换机的（这里的MAC是交换机背板的MAC），内容格式是：

Priority	MAC

Priority：缺省是32768，允许范围：1～65535
bridge－ID有两种
a) root-ID:记录根的，在BPDU的第一字段
b) sender-ID:记录该BPDU帧的发送者，即自己，在BPDU的第三个字段。

3. Port-cost:衡量一个接口所连线路速率的数值。代价值，越小越好。

带宽(单位：bit)	Port-cost缺省值
10G	2
1G	4
100M	19
10M	100

4. path-cost（线路代价）：一个交换机所连的网段到根所经过线路的port-cost之和。用来衡量一个交换机到根的距离。

STP实现的过程：
1. **选根**。
基于boot-ID字段，最小的是根。
（开机后，交换机都认为自己是根，所以root-ID的初始值是和sender-ID是一样的）
比boot-ID的大小，就是先比优先级priority,数值小的选中,priority相同的话,再比MAC值。数值小的选中
Priority的缺省值是：32768
确定根后，就只有根每隔2秒周期发BPDU了。非根交换机就只是转发BPDU了。
2. **选root port**
是为**非根交换机**选root port(root port:非根交换机离根最近的接口)。基于path cost选，小的优选。
3. **选指派接口**
为每个**网段**选一个指派接口（指派接口：网段离根最近的接口）。基于path cost选，小的优选。如path cost相同，比接口发的BPDU的sender-ID，小的选中.
4. 没被任何步骤选中的非根交换机接口，被该算法管理关掉，即被设为blocking.此时该接口只能接收BPDU帧，不转发BPDU帧，也不能收、发用户帧。

- 当交换机的root port失效时，被blocking的接口，会有一个被自动激活，成为active.成为root port(因为被blocking的接口能接收BPDU)
- 网络稳定后，接口只有2种状态：forwording、blocking.

接口的maxage:最大生存期，20秒。（BPDU在内存的最大逗留时间）也就是说原失效的root port相关的BPDU20秒后才会消亡，就是说20秒后才能确定root port失效了。

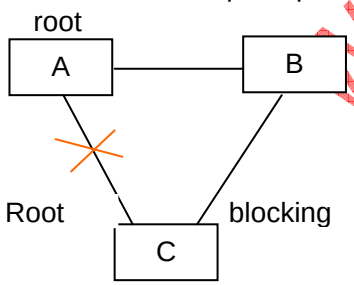
交换机接口的五种状态变化：

状态	从上一状态到该状态所需时间	从上一状态到该状态期间所做的操作	帧收、发状态		
			发BPDU	收BPDU	收发用户帧
Blocking阻塞	↓20秒 ↓15秒 ↓15秒	等待maxage消耗完 重新计算STP 交换机更新MAC表	×	OK	×
Listening监听			OK	OK	×
Learning学习			OK	OK	×
Forward转发			OK	OK	OK
disable	接口处于无插线状态		×	×	×

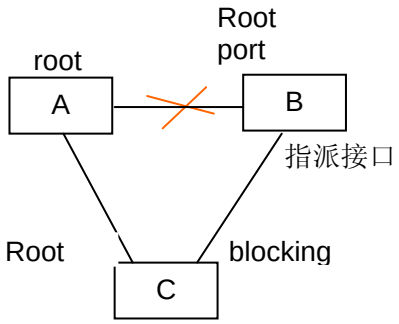
所以备份的blocking接口到起用（即forward），需50秒的收敛时间
接口插线到forward，即disable到forward，只需2个15秒，即30秒的收敛时间，不需要因为判断原root port是否失效而等待的20秒maxage。

Cisco用3个私有技术，加快STP收敛速度：

1. Port fast:（对PC客户端的）
   a) Port fast是针对接入层的用户端接口，能省掉2个15秒的延时（接口插线前disable到forward原需2个15秒），变成只需1～2秒。
   b) 设为port fast接口，只能接用户终端，不能接交换机。如接了交换机，当接口收到这个交换机的BPDU后，port fast会关闭（error disable），要进入到这个port fast接口做no shutdown打开它（因为接交换机，就不能保证不出现环）。
   c) 必须在准备连用户终端的每个交换机接口上去单独配置。
      如：int f0/1
          spanning-tree portfast
2. uplink fast（对自己的root port失效时，会改变自己的blocking接口的状态为forwording）
   a) 当一个非根交换机发现它的root port物理失效时，会立刻将blocking的接口立即转变为forwarding,省掉了50秒。
   b) C 的 blocking 接口跳变为 forwarding 时，会同时自动把自己的 MAC 表发给上游交换机。
      （因为当接口从 blocking 直接跳到 forwarding，没时间进行 spanning-tree 计算，A 根和 B 的 MAC 表没变化，只会等到该 MAC 条目 180 秒后自动清除）
   c) 针对一个交换机去配置，在全局模式下配，不用进接口模式。
      如：config t
          span uplink-fast



Uplink fast 图示



Backbone-fast 图示

3. backbone-fast（对没有blocking接口的交换机有接口失效时，会改变下游交换机的blocking接口的状态为forwording）
   a) 当一个无接口被blocking的交换机（这里是B）的root port失效时（通过**物理失败检测**，没被blocking的接口都是active的），它会通过发送TCN（通告BPDU）告诉其它非根交换机（这里是C）的，C会立刻发出帧询问其它交换机询问是否有到根的路径，得到**确定答复**后，C的blocking接口会跳过20秒的maxage(blocking到listending的时间)，经历2个15秒后到forwarding。**无确定答复**，C会重选root,重新技术STP.
   即，**一个树形结构有一处断了，一定会调整形状，重新满足树形通路。**
   b) 因为是互动的技术，所以要在所有交换机上配置 backbone-fast,（全局模式下）
   如：config t
   span backbone-fast

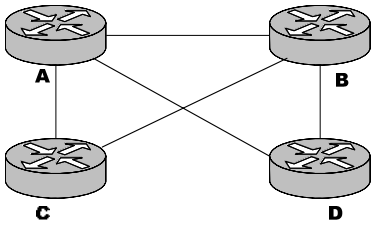
跟踪 STP 变化过程,用命令：
debug span

PVST:每个 VLAN 生成一棵树（cisco 私有的）
为了不浪费闲置带宽，在一定程度上实现 2 层的负载均衡。标准的 STP 技术 IEEE802.1Q 每个 VLAN 都用同样的线路，同样的 root port,不完善

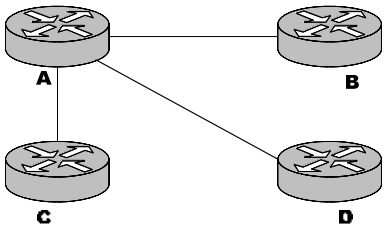
MSTP:（通用标准技术）把多个 VLAN 分成几个组，一组用一棵树。
如组 1：VLAN 1~ 500 用一棵树，组 2：VLAN 501~1000 用一棵树。
相对 PVST 而言，MSTP 在 VLAN 数很多时更实用。

RSTP:另外一个新的、标准的快速 STP 算法，cisco 没用，它有其它技术替代它。
改变 VLAN 22 的优先级为 4096，用命令：
spanning-tree VLAN 22 priority 4096
改变接口 port cost 为 18（是对所有 VLAN 都有效的），用命令：
spanning-tree cost 18
改变 VLAN 22 的所有接口的 port cost 为 17
spanning-tree VLAN 22 cost 17

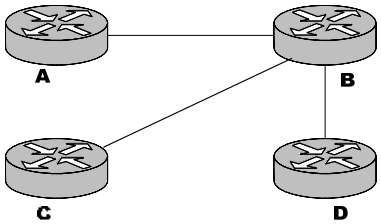
} 不提倡改 port cost,最好的方法是控制根的生成。



有 4 个 VLAN，用 2 棵树。
对于 VLAN1、VLAN2，A 做 root，B-C、A-D 线路不用。即所用线路为

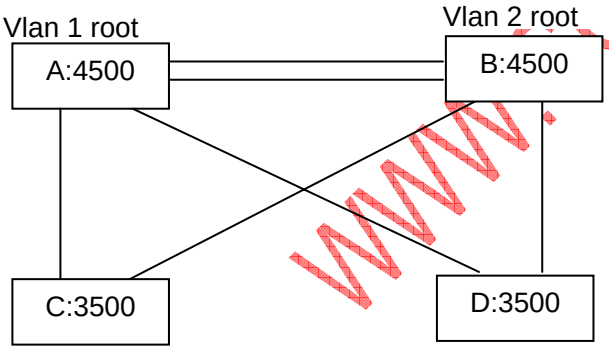


对于 VLAN3、VLAN4，B 做 root，A-C、B-D 线路不用。即所用线路为



A:spantree vlan 1 priority 10000　　(对于 vlan1 让 A 做根，须在 vlan1 中 A 的 priority 要最小)
　spantree vlan 2 priority 10000　　(对于 vlan2 让 A 做根，须在 vlan2 中 A 的 priority 要最小)
B:spantree vlan 1 priority 15000　　(对于 vlan1，让 B 下面的接口保证为 vlan1 的指派接口，不会成为 blocking，须将其 priority 设低于 32768 的缺省值)
　spantree vlan 1 priority 15000　　(对于 vlan2，让 B 下面的接口保证为 vlan2 的指派接口，不会成为 blocking，须将其 priority 设低于 32768 的缺省值)

B:spantree vlan 3 priority 10000　　(对于 vlan3 让 B 做根，须在 vlan3 中 B 的 priority 要最小)
　spantree vlan 4 priority 10000　　(对于 vlan4 让 B 做根，须在 vlan4 中 B 的 priority 要最小)
A:spantree vlan 3 priority 15000　　(对于 vlan3，让 A 下面的接口保证为 vlan3 的指派接口，不会成为 blocking，须将其 priority 设低于 32768 的缺省值)
　spantree vlan 4 priority 15000　　(对于 vlan4，让 A 下面的接口保证为 vlan4 的指派接口，不会成为 blocking，须将其 priority 设低于 32768 的缺省值)



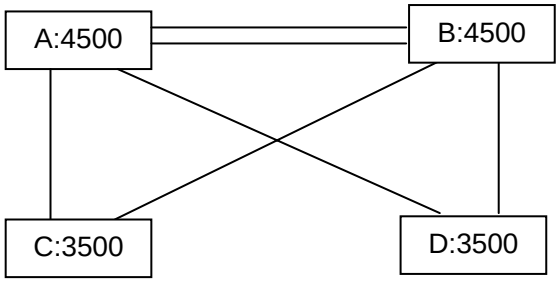
控制 STP 的配置操作：
用 PVST 为多个 VLAN 生成树时，要控制树的线路实现负载均衡，主要是通过**根的位置**来实现。不提倡改 port cost
A：Span vlan 1 priority 8000
B：Span vlan 1 priority 16000　} 把 A 设为 VLAN1 的根，B 下面接口为指派接口
A：Span vlan 2 priority 16000
B：Span vlan 2 priority 8000　} 把 B 设为 VLAN2 的根

在连电脑的接口上做 port fast，如：
C:int f0/1
　Span port fast

在 uplink 接口上做 up-link fast,如：（全局模式下）
C:span up-link fast

小工程上一般不做 backbone fast.

以太通道 Ethernet channel：2 个 core 核心层交换机之间用 2、4、8 根 trunk link 连接。为了高速传输且一定意义上负载均衡。
并且把这多个物理线路变为一条逻辑线路。避免了 STP 算法把一些端口 blocking 后，只有一条线路在用，其它做备份的情况。
注释：其它厂家有不同叫法,北电叫链路汇聚技术。



以太通道负载均衡实现：

1. 2 个物理线路分别编号为：0、1，用户发的帧的目的 MAC 地址的最后 1 位为 0 就走 0 号线路，为 1 就走 1 号线路。
2. 4 个物理线路分别编号为：00、01、10、11，用户发的帧的目的 MAC 地址的最后 2 位为 00 就走 00 号线路，为 01 就走 01 号线路，为 10 就走 10 号线路，为 11 就走 11 号线路。
3. 8 个物理线路分别编号为：000、001、010、011、100、101、110、111，用户发的帧的目的 MAC 地址的最后 3 位为 000 就走 000 号线路，为 001 就走 001 号线路，为 010 就走 010 号线路……
   注意：有时候，不一定用的是目的 MAC 地址，视厂家、型号而定，有时是源 MAC 地址，也可自定义。

以太通道连线规则：

1. 两边对应的以太接口，逻辑特性要一样。
   要么两边都是 trunk port(载波多个 VLAN)，要么都是 access port(载波 1 个 VLAN)
2. 两边对应的以太接口，物理特性要一样。
   a) 双工模式：
      单工：一边只能发，一边只能收；半双工：两边都可收发，同一时刻，一边只能发或收；全双工：两边可同时收发
   b) 速率
   c) 所有厂家设备以太接口的缺省值都是：**自动协商模式**，寻求两边物理特性相同的最高速连接方式。也可以不用自动，手动地设置，两边都设：（所有 cisco 的设备命令都一样）
      int f0/1
        speed 100　　(连接速率 100Mbit/s)
        dvplex full　　(全双工)
3. 在一边的交换机上，以太通道必须在同一芯片上最小号开始的连续的接口，否则会报错。
   交换机上的以太控制芯片：控制交换机的以太接口，一般 6 个接口用一个芯片（不是绝对的）。

PAGP 链路聚合协议：以太通道接口自动配置的协议（cisco 私有）
接口有几种状态模式：（缺省值是 desirable 渴望模式）

	ON	desirable缺省	Auto	Off
ON	生成	生成	生成	×
desirable缺省	生成	生成	生成	×
Auto	生成	生成	×	×
Off	×	×	×	×

也可手动配置：

```
int f0/1
  channel-group 1
int f0/2
```

channel-group 1
这样就生成了虚拟的 PAGP 组接口：channel-group 1 ，对这个组接口发布命令，会对组内所有的实际接口都起作用。

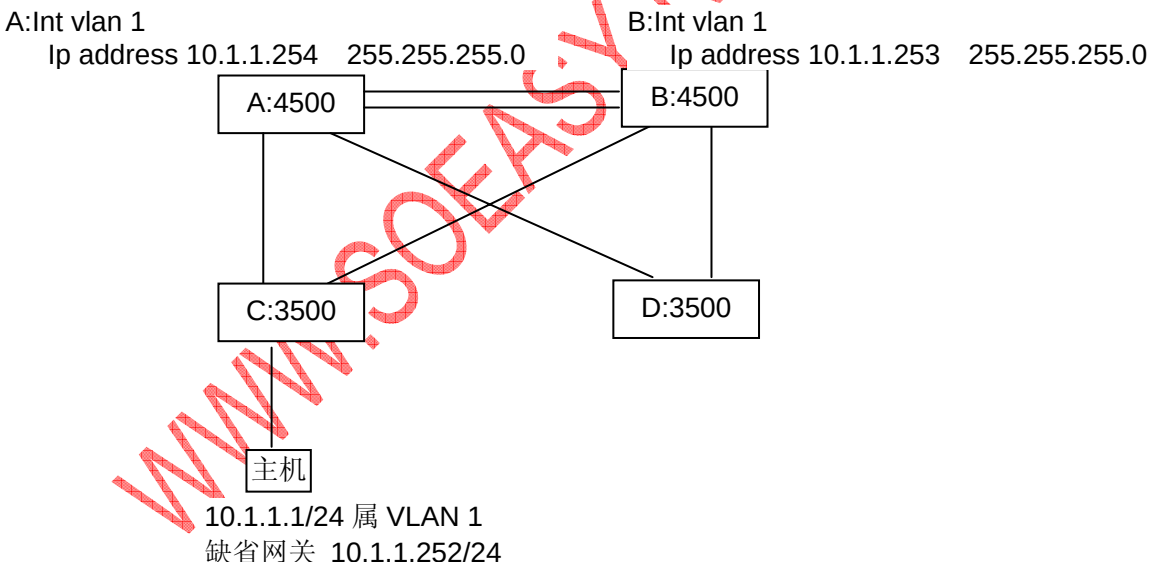
假如两个没配置过的 4500，用 2 条交叉线连，PAGP 会先检测运行，DTP 会检测运行。

冗余：准备概念：
实际上上面说的 SPT 是 在拓扑上做了冗余。
cisco 4500、6500 是模块化的设备，刚买的设备只有一个主板，电源、模块都要另外买。(MSFC 卡相当于一个小路由器？)
超级引擎模块：上面有 CPU、内存、flash、ios 操作系统等

1. 超级引擎冗余：4500、6500 可配 2 个超级引擎，且只能插在 1、2 槽上，1 槽运行、2 槽备用（只有一个引擎时就只能插 1 槽）。2 个超级引擎之间的冗余控制协议是：
   RPR（缺省值）：比较早期的，备用的引擎没运行，切换生效需 2～4 分钟完成。
   RPR＋：备用的引擎同时运行，且内存同步，切换只需 30 秒～1 分钟。
   可手动改缺省值，改为 RPR＋：redundancy (冗余的意思)
   mode rpr-plus
2. 电源冗余：4500、6500 可配双电源，
   缺省是只开一个（冗余的 redundancy）,还可以一起用（组合的 combined）
   power redundancy-mode redundancy
   power redundancy-mode combined

HSRP 热备份主机协议（cisco 私有的）：做缺省网关冗余，即做 2 个缺省网关，以策备用。（原有的代理 ARP 不好）

提问：为什么要缺省网关？
回答：VLAN 间要路由。网内的主机要路由就要找到路由器，路由器的接口就是这个 VLAN 的缺省网关。



上图，A、B 两个可以做路由的 3 层交换机在 VLAN1 子网内都设了可路由的接口 10.1.1.254/24、10.1.1.253/24（都可单独做这个子网的缺省网关用），用 HSRP 把两个接口虚拟到单个缺省网关上 10.1.1.252。
同样，A、B 两个也可以在 VLAN2 子网内都设可路由的接口 10.1.2.254/24、10.1.2.253/24，用 HSRP 把两个接口虚拟到单个缺省网关上 10.1.2.252。
VLAN1 中的两个真网关，其中一个是备份的状态（stundby）,另一个是激活状态（active）,由它来响应主机们发往虚拟网关 10.1.1.252 的 ARP 请求。

HSRP 组 ID:取值范围是 1～255
（在一个 VLAN 里如果有，就只能有一个 HSRP 组，因为一个 VLAN 只有一个网关。同是一个 HSRP 组内的接口，才能协同 HSRP 操作。）

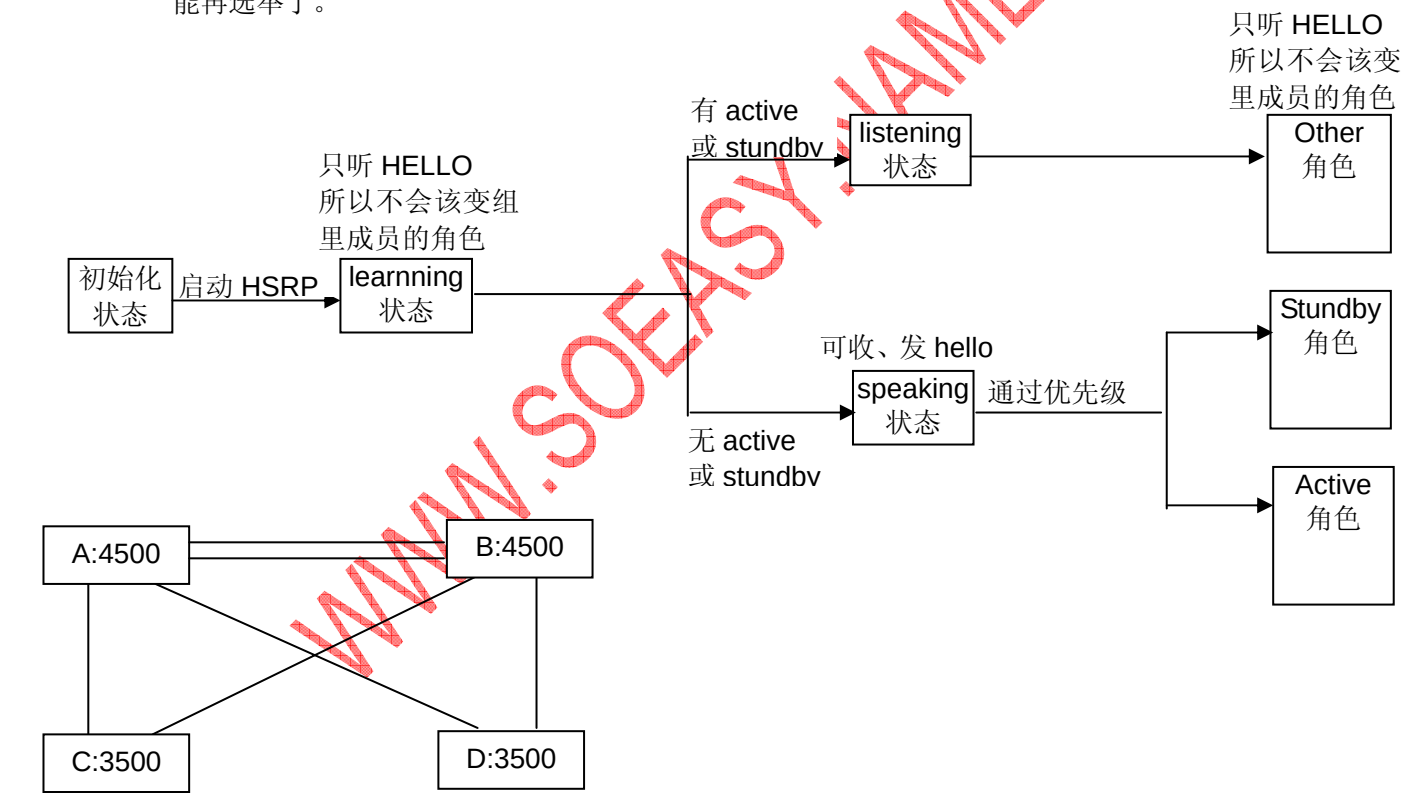
参与 HSRP 操作的接口的角色：（必须是有能力通过一些办法实现路由的接口，才有资格参加 HSRP）

i. active：组成虚拟网关的两个接口之一，响应用户发往虚拟网关的 ARP 请求的那个接口
ii. stundby：组成虚拟网关的两个接口之一，不响应用户发往虚拟网关的 ARP 请求的那个接口
iii. other：组中除上述两个接口之外的接口。

HSRP 的选举过程：

概念：

- 选举基于 HSRP 优先级，其取值范围 1～255，缺省值是 100，高的优选。其次必接口 IP 地址，高的优选。
- 同一 HSRP 组内接口之间，每隔 3 秒周期性的用 Hello 包交换 HSRP 信息（包括各自知道的接口的优先级等）
- holdtime:缺省 10 秒。当 10 秒内，stundby 一直从 hello 信息中都没得到 active 还存在的消息时，就会转变为 active 角色。（如果老的 active 恢复了，虽然它的优先级高，但会由于选举过程步骤原因而不会夺回 active 的角色，是为了网络的稳定。除非这个接口上面配置了占先指令。如 standby 1 preempt.）
- active 在第一次是选举确定的，之后如果该 active 失效了，再都是从 stundby 切换过来的了。不可能再选举了。



HSRP 的配置指令：

做 VLAN 1 的 HSRP 组：

```
A:int vlan 1
    Ip address 10.1.1.252   255.255.255.0
    Stundby 1 ip 10.1.1.254                 （在 A 的 VLAN1 上指定 HSRP 组 1 的虚拟网关为 10.1.1.254）
B:int vlan 1
    Ip address 10.1.1.253   255.255.255.0
    Stundby 1 ip 10.1.1.254                 （在 B 的 VLAN1 指定 HSRP 组 1 的虚拟网关为 10.1.1.254）
A:int vlan 1
    Stundby 1 priority 120                  (把优先级从缺省的 100 变为 120，使 vlan1 的根：A 成为 active 角色)
```

Stundby 1 preempt （占先指令，A 失效又恢复后，会夺回 active 角色的。见 HSRP 概念 3）

做 VLAN 2 的 HSRP 组：

A:int vlan 2
Ip address 10.1.2.252 255.255.255.0
Stundby 2 ip 10.1.2.254 （在 A 的 VLAN2 上指定 HSRP 组 1 的虚拟网关为 10.1.2.254）

B:int vlan 2
Ip address 10.1.2.253 255.255.255.0
Stundby 2 ip 10.1.2.254 （在 B 的 VLAN2 指定 HSRP 组 1 的虚拟网关为 10.1.2.254）

B:int vlan 2
Stundby 2 priority 120 (把 B 优先级从缺省的 100 变为 120，使 vlan2 的根：B 成为 active 角色)

看 HSRP 接口状况变化过程：debug stundby
看 HSRP 接口状况列表：debug stundby

HSRP 接口跟踪（tracking）技术：当一个 VLAN 的 active、stundby 都连接到远程网络时，active 失效了，stundby 可以和它对调角色的技术。保证到远端网络的传输。也是通过对优先级的操作实现的。

多层交换：

交换一般用硬件完成查询 MAC 表来处理帧。

路由一般用软件完成查询路由表来处理 IP 包。

硬件：（ASIC：大规模集成电路芯片。）

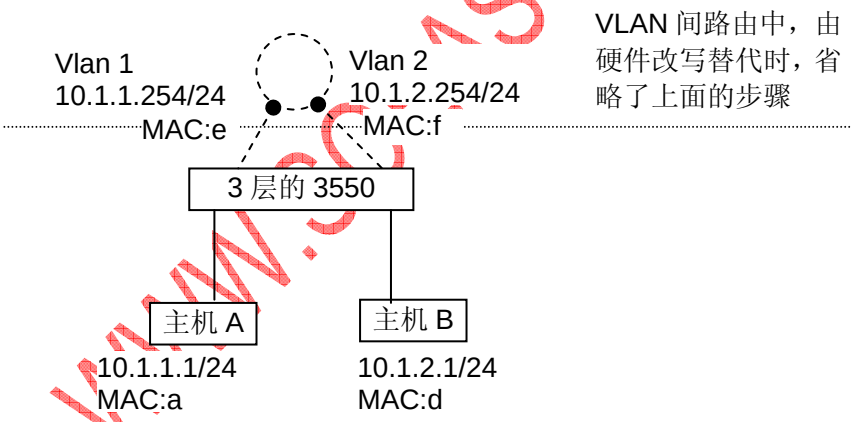
CAM 芯片：存放 MAC 表（也可以是路由表等数据表格，一般情况路由表是放在 RAM 中的）。可自主查表。

TCAM 芯片：CAM 的第二代。

用 2 层的硬件完成 3 层查询路由操作的技术：

1. MLS 技术（mutli layer swith）：针对流的操作去硬件执行。对 TCP 会话的不同，如 ftp、http 分别处理操作。（具体略）
2. Cef 技术：针对网络拓扑去硬件执行。对路由中改变帧和帧内 IP 包的内容操作不同而分类，去分别用硬件替代执行。执行结果是不变的。
   a) Fib 表（转发分类信息库）：记录了那类包可以做那类硬件操作的分类信息，且与路由表自动同步变化，实际上，fib 表是路由表的简化版。
      i. Fib 表是从路由表自动推算出来的，而且是存放在 TCAM 中，由 TCAM 自主查询不需 CPU 参与的。路由表是路由器内存，由路由器 CPU 查询的，所以 fib 的硬件操作比路由表的软件查询快的多。
      ii. Cef 也可以达到负载均衡的目的，因为 fib 表是从路由表推算出来的。而且硬件改写与路由表操作结果一样。
   b) 调节表(adjacency)：记录了 fib 表中对应类的具体操作方式。由 ARP 表推出的，由指针来实现与 fib 表对应。

      是从第一次的真实路由过程中的真实 ARP 操作自动学习来的。以后只是用硬件执行有变化的部分。



主机 A 发出的帧：

源 MAC:a	目的 MAC:e	源 IP：10.1.1.1	目的 IP：10.1.2.1	ICMP 请求

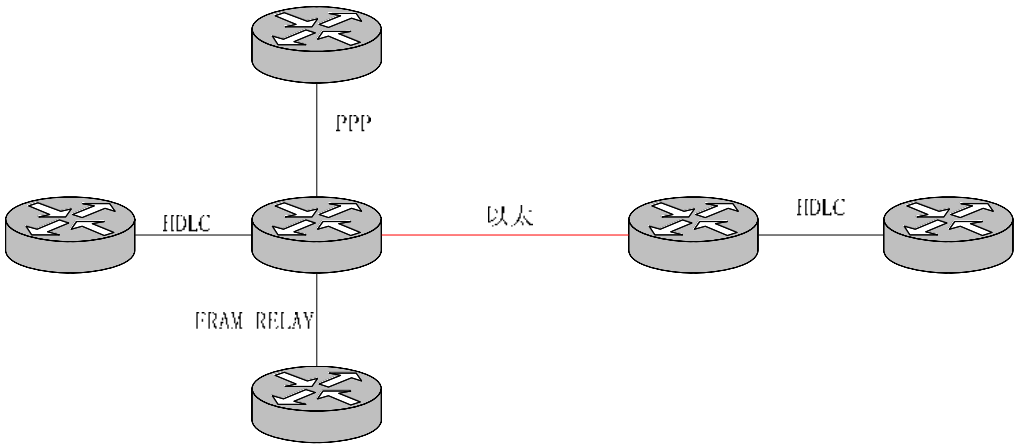
主机 B 发出的帧

源 MAC:c	目的 MAC:f	源 IP：10.1.1.1	目的 IP：10.1.2.1	ICMP 请求

其中帧有变化的字段：源 MAC、目的 MAC、FCS(帧校验字段)

其中 IP 包有变化的字段：TTL 字段－1（路由的次数的倒计时）、check sum(IP 包校验字段)

在 cef 算法的控制下，这些都被直接硬件改写。



激活 cef 的指令：
ip cef (缺省值)
关闭 cef 的指令：
no ip cef
查看 fib 表：
show ip cef
查看调节表：
show adjacency

组播：局域网、广域网的视频点播。广域网中设备太乱，所以没实际实用。

组播 IP 地址是在第三方软件中设置的，不是在操作系统中设置，操作系统中只能设置单播地址。

1. IP 组播地址：多个单播 IP 地址组成的组的号码。

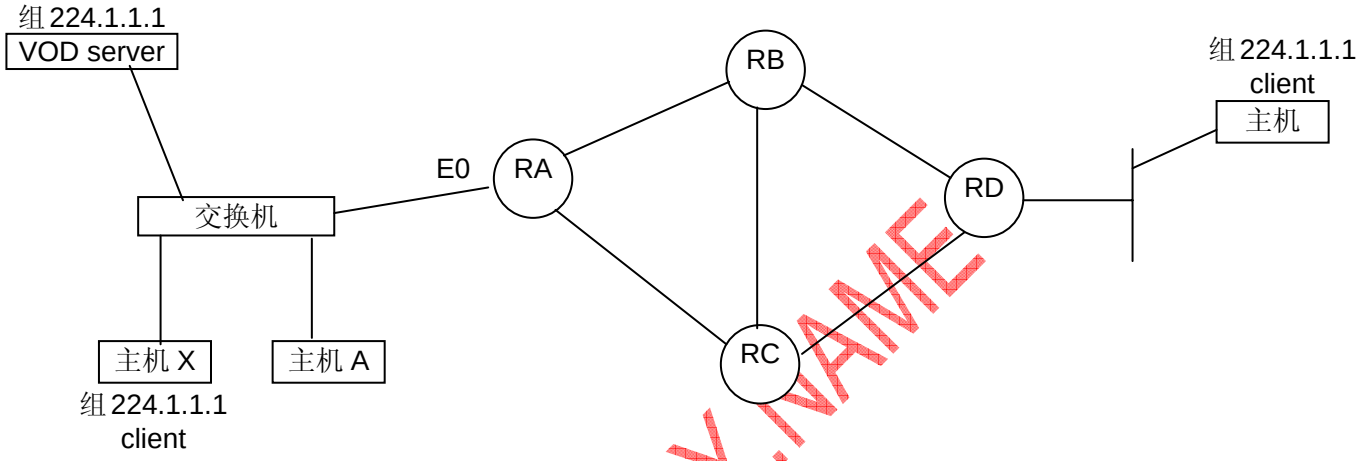
   组播地址：D 类：224.0.0.0～239.255.255.255 无掩码！

   单播地址：A~C 类

   广播地址：主机位全为 1

   保留地址：E 类

2. MAC 组播地址： 组播的以太地址，8 个组播 IP 地址用某算法转为 1 个 MAC 地址 等等。



1. 子网内组播：

   VOD server 发的组播包，经封装为帧发到交换机，一个透明桥接的交换机会从所有接口转发出去。

   VOD server 发出的帧的

源 IP	目的 IP	源 MAC	目的 MAC
接口 IP	组 IP：224.1.1.1	接口 MAC	虚拟的，计算的

2. 子网间的组播：

   路由器接口缺省模式是丢弃帧广播、组播信息。

   打开路由器接口接收广播帧、组播帧指令：

   config t

   ip mulit-costing-routing

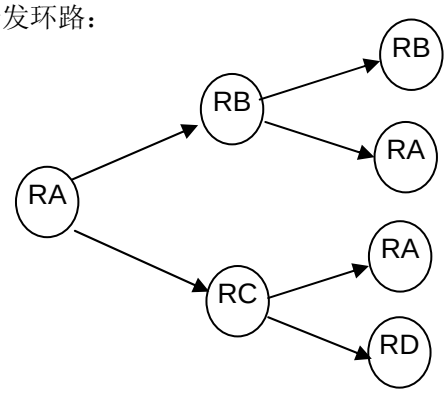
   子网间的组播的两种方式：（在组播前，路由器要做好单播路由表）

   1 soure-base-tree:在路由网络模型中，每个组播源路由器（如 RA），以自己为根，生成以到每个组播目的路由器（如 RD）最短路径组成的树，（有几个源就有几个树）。

   对应协议为 PIM-DM:是密集模型，适用组播接收者多，即接受者高密度，是带宽要求较高的方式。

   PIM-DM 协议：源路由器从所有激活了 PIM-DM 的接口转发组播包（除了接收组播的接口<如 RA 的 E0>以外。）

   出现组播转发环路：



用 RPF（反向路径检测）技术解决：组播接收路由器，比较到组播 IP 包中源地址的最近路径的出发接口（即路由条目的相关接口）与接收接口是否是同一个接口，如是才接收这个包。否则，不从这个接口接收这个包。

2 Share-tree:所有源用一棵树。设定了 RP 汇聚点路由器,所有源都把 IP 包发给 RP，由 RP 发往目的路由器。

对应协议为 PIM-SM:是稀疏模型，适用组播接收者少，即接受者低密度，是带宽要求低的协议。

PIM-SM:RP 精确组播，组播接收路由器（如 RC）会告诉 RP（汇聚点路由器）它接收的哪个组的组播，RP 就精确发给它对应组的组播包。

IGMP(internet group manage protocol):路由器管理所连子网中的组的管理，包括组的删增改。

CGMP（cisco group manage protocol）:解决如主机 A 这样的电脑也要接收判别不相关组播包的现象等问题（浪费主机资源、网络带宽）的管理协商协议。

RA 路由器收 IGMP 信息，处理组的数目及其成员信息对应的 MAC 关系，转成 CGMP 信息发给交换机，交换机就知道哪些接口参与了哪些组的组播。

组播的配置：（当然要先配置好单播路由表）

PIM-DM 组播的配置：

Ip mulit-costing-routing (开组播)

在每个要想参与 pim-dm 的接口上开 pim-dm，如：

int s0

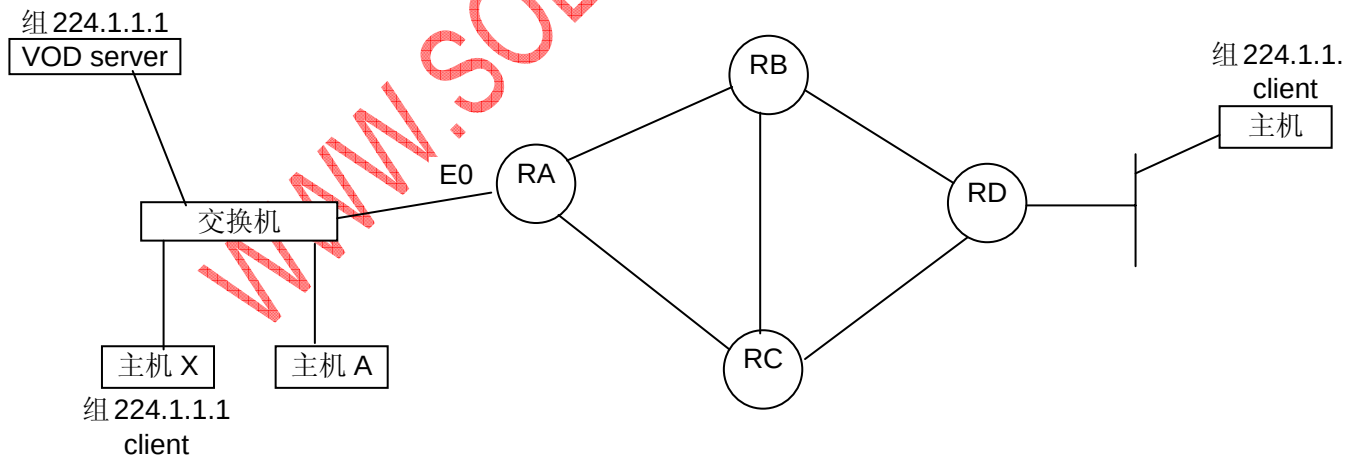
ip pim dense-mode

测试组播成功与否的方法:

RC:int e0

Ip igmp join-group 224.1.1.1 (把 RC 的 e0 设为 224.1.1.1 组的接收者，IGMP 默认关闭)

RA:ping 224.1.1.1 (用 ICMP 的 PING 来模拟发出组播包的出发路由器)



PIM-SM 组播的配置：

Ip mulit-costing-routing (开组播)

在每个要想参与 pim-sm 的接口上开 pim-sm，如：

int s0

ip pim sparse-mode

生成 RP（在所有参与 pim-sm 的路由器上都执行这个命令）

ip pim send-rp- announce Lo0 16 （Lo0 是 RB 的 lookback 接口，16 是限制的最大跳数，由 TTL 字段记录）

测试组播成功与否的方法:

RC:int e0
　　Ip igmp join-group 224.1.1.1　　(把 RC 的 e0 设为 224.1.1.1 组的接收者，IGMP 默认关闭)
RA:ping 224.1.1.1　　(用 ICMP 的 PING 来模拟发出组播包的出发路由器)

用到的相关命令：
查看参与这个 pim 的接口有哪些
show ip pim neihbor
查看组播路由表
Show ip mroute

远程访问：

概念：

- 广域网属于是通信服务供应商（如电信等）提供的线路。
- 在两个城市间铺有 2Gb/s 至 40Gb/s 左右的光纤，两端是 SDH 设备（光传输终端设备）。通过时分多路复用，每个时隙（时间片）提供 64Kb/S 的带宽。
- 包交换网络（packet switch network）：在城市间光纤两端，另外有成对包交换机。包交换机上都连着多个用户，这些用户共用给予包交换机的带宽（共享着若干个时隙）。有可能 2Mb/s 的包交换机线路被几十个用户在同时共用。

- 广域网的 piont-to-piont 网络类型的帧封装协议主要有：PPP、HDLC（局域网帧封装协议：ethernet 2、IEEE802.3、IEEE802.3-SAP、IEEE802.3-SNAP）
- 几种通信协议的带宽：E1=2.048Mb/s（30 个时隙）、T1=1.544Mb/s（24 个时隙）、E3=155Mb/s、T3 大概 38.6Mb/s，E1、E3 用于欧洲和中国等，T1、T3 用于北美和日本等。E3、T3 一般用于 ISP 商之间的连接。
- CISCO 路由器分类：
  - 中心级别（一般在总公司用）：cisco10000(数千万/台)，7500（数百万/台）7200，3700，3600，2600 系列。
  - 分支级别（一般在分公司用）：3700，3600，2600，1700（5000 元左右）
  - 办事处级别：cisco1600
  - 家庭办公：cisco800

广域网线路租用分类：

1. DDN 专线(也叫租用分类、2M 透传等)：适用于各地分公司的主干连接。点到点的，只提供物理层（1 层）的通道。
   - 选择速率的范围大：从 9600b/s 到 Gb/s 级的任意选择。
   - 仅是租用者单独使用，其它人不可能使用，所以安全（你使用的时隙，你不用时会空着，不会给别人用）。
   - 因为是租用者单独使用，所以拥塞机会极小，稳定。
   - 永久连通。
   - 缺点：费用高
2. 拨号线路（也叫电路交换。如：PSTN 电话系统，DSL 综合数字通讯）：适用于各地分公司的备份连接。点到点的，只提供物理层（1 层）的通道。
   - 平时不连通，需传输时连通
   - 成本低
   - 缺点：速率低。
3. 包交换网络（packet switch network，也叫分组交换）:点到多点的，提供了物理层、链接层（到 2 层）的通道（因为提供了帧的封装格式：ietf、cisco）
   - 选择速率的范围大
   - 成本相对较低，大概是 DDN 的一半费用。
   - 永久通道
   - 所用技术：
     - 帧中继。（现在的主要的包交换技术）
     - ATM（异步传输模式），已被淘汰
     - X.25，已被淘汰

- 缺点：可能会随机产生拥塞。（因为是多用户共用）
- 缺点：有安全隐患。（因为是多用户共用）

远程连接的网络类型
point-to-point 网络类型：DDN，拨号线路
point-to-mutlipoint 网络类型:包交换网络

AAA 体系（authentication authorization accounting 认证 授权 统计）：在路由器上用的
认证：确认用户身份是否合法
授权：确定用户可以作那些操作
统计：确定用户作了哪些操作

RAS 远程访问服务 remote access services：（属于拨号接入的）
RAS 服务器：远端用户经 PSTN 或 ISDN 等拨号连接 RAS 服务器
RAS 服务器设备的选用：

- 专用设备：如华为 8010(700 万/台)，CISCO 5800、CISCO 5300 等
- 经特殊配置的路由器：如 CISCO 2600、3600 可配置层 RAS 服务器
- 普通电脑：以 NT、2000server 为操作系统，打开用户接入的功能

（cisco2500 是固定接口，cisco2600、3600 是模块化的设备，有 8AM、16AM、32AM 的 modem 卡模块来配置，能使其具有 8、16、32 个 modem 的接入口）

用户认证的协议(PPP 协议中的子协议)：

- CHAP：用户传往 RAS 服务器的帐户及密码信息被加密、用散列算法来保证完整的认证过程协议
- PAP：用户传往 RAS 服务器的帐户及密码信息不被加密、不用散列算法来保证完整的认证过程协议

ACS 软件（当前最高版本是 3.0）：CISCO 私有的，而且在授权、统计方面不怎么好，有另外公司编写的软件比较好。

- 在有多个 RAS 服务器时，用运行 ACS 软件的电脑来集中查询认证，而不用在每个 RAS 服务器上建立拨号用户的帐户及密码信息的数据库。
  PAS 服务器向 ACS 运行服务器发送用户帐号密码信息去验证时也加密，所用协议为：（这个加密过程是在 7 层即应用层加密的）
  - TACACS+协议（TCP 传输）cisco 私有的：把帧中所有的数据部分全部加密
  - RADIUS 协议（UDP 传输）通用的：只对帧中的数据部分的用户帐户密码信息加密。
- 另外，可以将这个数据库放在一个数据库服务器（如 oracle、DBR 等）上，让其它应用也调用该数据库。

CISCO 的 AAA 认证流程的角色：
AAA user:远程访问者
AAA client:RAS 服务器
AAA server:ACS 服务器（或在 RAS 服务器本地验证）
AAA database:数据库服务器（或在 ACS 服务器上、或 RAS 服务器本地）

同步：靠时钟来区分字节开始和结束的位置。如 serial 0 接口。
异步：靠起始符和停止符来区分字节开始和结束的位置（效率低，因为要传送非数据的起始符）
对于广域网的 piont-to-piont 网络类型的帧封装协议：PPP、HDLC

远程连接 2 层封装协议的通讯模型：
同步：PPP 、HDLC
异步：PPP

远程连接不同方式的通讯模型：
同步：（数字线路）DDN，包交换
异步：（模拟线路）拨号线路

远程拨号的典型网络模型的 3 个层：

- 物理层：PSTN、ISDN 线路
- 物理链路层：PPP 协议
- 网络层：IP 协议

AAA 服务的认证服务的用户身份确认针对 2 种访问模式的用户：

- 包模式（packet 模式）：要访问网内资源的用户的访问方式。会认证物理层、物理链路层、网络层共 3 层的连接。
- 字符流模式：要配置设备，进入 IOS 命令行的用户的访问方式。只认证物理层的连接。

（命令 line:配置的是物理层的；命令 interface：配置的是物理链路层、网络层的。）
如：line 1 8　(配置 MODEM 模块卡 8AM 的 8 个口)
line con 0
line aux 0
line vty 0 4

配置 CISCO 的 AAA 的认证（这里没配授权、统计）
aaa new-mode　(打开 aaa 服务，路由器缺省是关闭的。)
该命令发出后会产生个隐形命令发出
aaa auth login defult local　defult 是缺省的认证方式，即本地认证
login 是字符流模式的认证模式，也可以是 ppp 等包模式的认证方式。
Line vty 0 4　(虚拟终端登陆认证)
该命令发出后会产生个隐形命令发出
aaa auth defult　defult 是缺省的认证方式，即本地认证

Line 下字符流模式
username xxx password xxx　建立本地认证数据库
aaa auth login xxx　XXX 认证方法的名字
group tacacs+
tacacs-host ip 10.1.1.1/24　定义认证服务器的 IP 位置
tacacs-host key cisco

标准 HDLC 帧的字段组成：
flag:帧起始符
address:因为是点到点的，所以这个地址有无都无所谓，这是个虚假的单地址
control:
data:
fcs:帧校验和
flag:帧结束符
HDLC 没有说明上层协议类型的字段，不能同时支持多种 3 层网络协议（即不能支持上层的多网络协议的多路复用），如 IP、IPX 等
CISCO 的 HDLC 帧：（私有的）
加了个网络类型字段。所以支持上层多网络协议的多路复用。
华为、CISCO 的同步接口（就是同步口 S0、S1 等）缺省的封装模式是：HDLC

PPP 帧的字段组成：加了 NCP、LCP 字段
PPP 的特点：

1. 支持同步、异步
2. 支持上层的（即 3 层的）网络协议的多路复用
3. 支持多种功能：(LCP 来实现)
   a) 身份认证
   b) 多链路绑定（类似以太通道技术）
   c) 错误检测恢复（重发）
   d) 回拨 call back（用低费用方式回拨远程访问者，类似手机接到电话就挂断，用小灵通回拨。）

ppp 网络模型:

网络层		
链路层	NCP	LCP
物理层		

NCP 网络控制协议 network control protocol：用来支持上层（即 3 层的）网络协议的多路复用
LCP 链路控制协议 link control protocol：实现 PPP 的多种功能。

DDN 的配置
配置顺序：1 连通物理层 2 物理链路层（缺省是 HDLC） 3 网络层，即 IP

1 连通物理层：即激活
2 链路层：
A：和 B: int s0
Ip address ……
No shutdown (查看：show int s0)
Encap ppp (物理链路层从缺省的 HDLC 改为 ppp，两端都要改)

(ppp，认证方式定为 chap 加密)
A:int s0
Ppp auth chap (认证方式要设的话，A、B 都要设，且要两边对应)
Hostname a
Username B password cisco hostname A
B:int s0
Ppp auth chap
Hostname b
Username A password cisco hostname B

接口的类型：
RS232:：主要用于异步通讯模式，如 console 口等
V.35：主要用于异步通讯模式
DB60（cisco 私有的，60 芯）:主要用于同步通讯模式，如 serial 口

通讯的协调：
DTE(data terminal equipment)类型
DCE（data communications equipment）类型

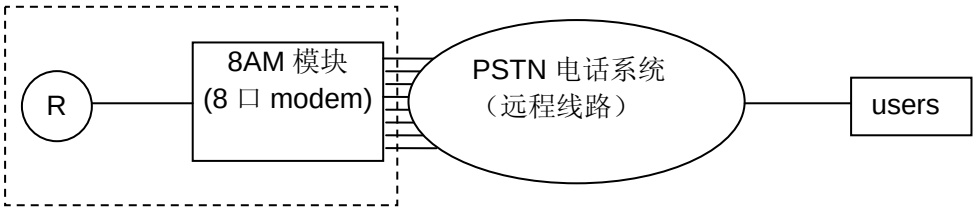
同步：DCE:时钟的发出者
DTE:时钟的协调者
异步：DCE:modem
DTE:路由源

流控：硬流控：通过引脚
软流控：通过窗口大小

关于 modem:
modem 在超级终端的命令行命令：
AT&F：把 modem 恢复为出厂值
ATS0=n：n 分钟后应答连通。

在 cisco 路由器上的 modem 模块卡上装 modem 驱动。（这是在 modem 1 到 8 路上配）
line 1 8
modem autoconfig discovery (即插即用的方式装 modem 驱动)
show modem cap (查看 modem 驱动库中已有型号的驱动。)
如果无该 modem 型号驱动,S0 会用 defult 驱动使用。
Modem cap edit (在 modem 驱动库添加驱动程序)

用 modem 作远程接入服务的配置：



字符流模式：

```
 line 1 8                          （也可以在 AUX 0 作远程拨号，对于外地用户，用远程拨号配置，可以人
                                   不到现场，在办公室就可以调错。）
    modem autoconfig discovery     (即插即用的方式装 modem 驱动)
    modem dial-in                  (允许自动应答)
```

包模式：

```
 line 1 8
    modem autoconfig discovery     (即插即用的方式装 modem 驱动)        ┐
    modem dial-in                  (允许自动应答)                       ├ 物理层
    ppp autoselect                 (ppp 从缺省只支持字符流模式，也支持包模式了) ┘
    interface async-group 1        ┐
      group-rang 1 8               ┘把 8 个 modem 口合为一组一起配，组名为 1
        encap ppp                  (PPP 封装)                           ┐
        ppp auth chap              (认证加密)                           ├ 链路层
        async-modem dedicate       (专属，只允许包模式)                 ┘
        username cisco password cisco (建立认证数据库 cisco)            ┐
        ip unnumber e0             (8 个 modem 口都没自己的 IP，都用 E0 的 IP 地址) ├ 网络层
        peer defult ip pool cisco  (给用户分配的 IP 地址池名称叫 cisco) ┘
    username cisco password cisco
    ip local pool cisco 10.1.1.2   10.1.1.10 (cisco 的地址池范围为 10.1.1.2   10.1.1.10，其掩码用 E0 的填充)
```

在远端路由器上设置了允许字符流模式 modem 拨入后，即在“超级终端”输入：ATDT（拨号），就建立了远程连接(相当于用 2 个 modem 很大的延长了串口线。)，就可以登陆远端路由器的命令行 CLI。(另外如果在 win2000 中，在“网上邻居”，“建立连接”，选“接受传入的连接”，也可允许远端 modem 拨入。)

ISDN 服务根据不同的信道数目分为：

1. BRI(Basi Rate Interface)：典型的 ISDN　BRI 服务提供两个 B 信道和一个 D 信道　（2B+D）
2. 主速率接口 PRI（Primary Rate Interface）：
   a) [北美和日本]提供的 ISDN　PRI 服务提供二十三个 B 信道和一个 D 信道（23B+D）速率可达 1.544Mbps ＝T1
   b) [欧洲]提供的 ISDN　PRI 服务提供三十个 B 信道和一个 D 信道　（30B+D）速率可达 2.048Mbps

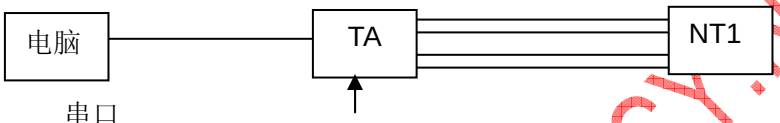
ISDN 综合业务数据网（一线通）：本来是用来替代 PSTN 的模拟电话网的，它是数字信号的。

逻辑：4 芯 S/T 接口　逻辑：4 芯 S/T 接口　逻辑：2 芯 U 接口
外形：RJ-45（8 芯）　外形：RJ-45（8 芯）　外形：RJ-11（4 芯）

ISDN 设备的说明：

TE1:具有 S/T 接口，并且可识别 ISDN 信令（如 TE1 的电话。如果 CISCO 路由器插了 wic-2b 卡(具有 2 个 bri 接口)，就有了 4 芯的 S/T 接口，使路由器也属于了 TE1,可直接连接 NT1.）

TE2:虽然有 S/T 接口，但自身不能识别 ISDN 信令（如电脑）



NT1：不用电源，由 ISDN 网络供电，就象 PSTN 模拟电话网一样。

ISDN 的物理层意义：2B+D 。B：64kb/s,可作数字或语音传输，在不同一时刻。D：16kb/s,信令通道，控制连接用的。总带宽 144kb/s,有 128kb/s 传数据。

在 cisco 路由器上驱动 ISDN 的指令：在全局配置模式下，ISDN switch-type base-net 3　(base-net 3 是 ISDN 的其中一种标准)

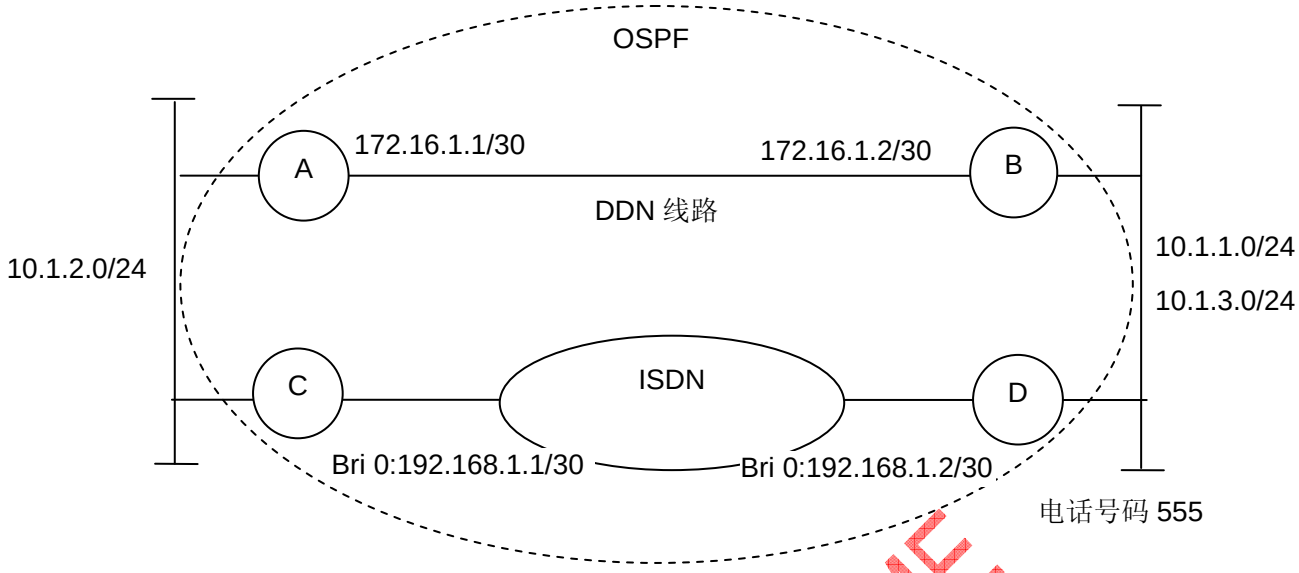
ISDN　BRI 需要配置：

1. 交换机类型：中国是 isdn switch-type basic-net3
2. 服务提供者标识符：　(电话号码，用户，口令)
3. 进行动态拨号配置，也可以使用 NAT 上网功能

DDR:按需拨号系统。

用 ISDN 线路作 DDN 的备份线路，目的：当 DDN 这个主干 down 了，ISDN 就拨通；当 DDN 这个主干恢复了，ISDN 就断开

有趣传输：用访问列表定义什么样的包可触发接口拨号。



先做 RAS 服务器:

```
D:isdn switch-type base-net 3                    （驱动 ISDN,为 base-net 3 标准）
  Int bri 0
    Encap ppp                                    (链路层的封装定义为 PPP)
    Ppp auth chap
    Ip address 192.168.1.2 255.255.255.252
    Hostname D
    Username C pass cisco                        } 与 C 的对应
```

```
C:isdn switch-type base-net 3                    (驱动 ISDN,为 base-net 3 标准,与 D 对应）
  Int bri 0
    Encap ppp
    Ppp auth chap
    Ip address 192.168.1.1 255.255.255.252
    Dial-group 1                                 (在 bri 0 上调用 dial 组 1)
    Dial map ip 192.168.1.2 name D 555 broadcast (电话号码 555，并且允许广播包)
    Dial time-out 20                             (线路空余 20 秒后，自动断开)
  Router ospf
    Redis  static  router-map                    (在 C 上做静态到 OSPF 的重分布，让 C 用 OSPF 告诉 A，在 B 那边有哪些网段。因为，AB 间的线路断了，A 已不能从 B 那通过 OSPF 知道这些网段)

  Hostname C
  Username D pass cisco                          } 与 D 的对应
  Access-list  1 permit 10.1.2.0  0.0.0.255
  Access-list  1 protocol ip list  1             （把 list 1 的包变为有趣传输）
  Ip route 10.1.1.0  255.255.255.0  bri 0 250   (静态的下一跳改为 bri 0 接口【必须为点到点】，浮动静态：把静态的管理距离变为大于 OSPF 的 250，当 A 的 OSPF 挂了，静态的这条路由条目就可进入 C 的路由表，C 就拨号了。)
```

A、C 这边的网段 10.1.2.0/24 里的机器的网关一般指向 A。

当右边的 10.1.1.0/24 自己 down 了，C 关于这个网段的静态路由进入 C 的路由表，C 就拨号了。但这个时候 DDN 是好的。C 发包，而 D 把到 10.1.1.0/24 的包丢了。

可以用这些语句替代 DDR：(把上面 3 行改为下面 2 行)

```
C:dial-watch-list 1 ip 10.1.1.0                  } 当这些条目都没了，就是右边的网段 A 一个都去不了的时候，C 才拨号
```

dial-watch-list 1 ip 10.1.3.0

包交换网络（packet swith network）:（类似以太网，只是距离很长）

包交换网络本身就定义了帧：帧中继帧，一：ietf 帧中继帧，二：cisco 帧中继帧

VC:一条虚拟链路，在 2 个通讯终端之间进行虚拟连接。

一个物理接口可承载多个 VC。

DLCI:唯一标识一条 VC，它具有本地有效性。它是一个数字，由电信局分配的，我们不能改变的。

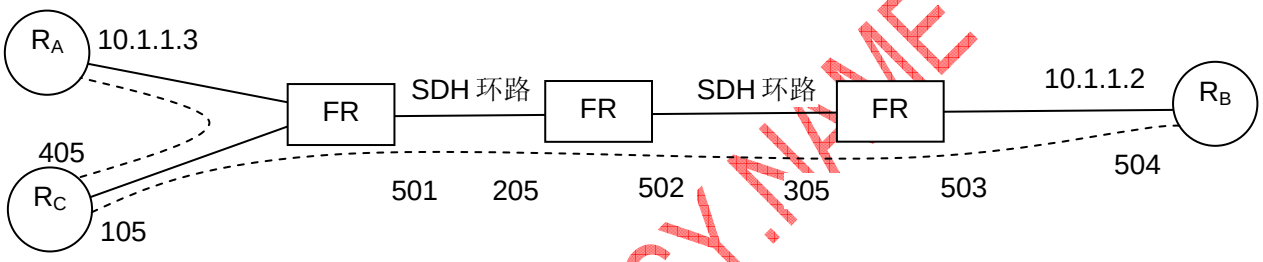
LMI（local manage interface）:运行在路由器、交换机之间，它协商 VC 的建立过程。它来协商帧中继连接的参数。

LMI 的 3 个标准：

- CISCO（默认值），自适应方式的。可自动检测对方的标准，去改变去适应。
- ITU-T，国际电信联盟。参数是:**Q.933A**
- ANSI，美国国家电器协会。参数是:**ANSI**

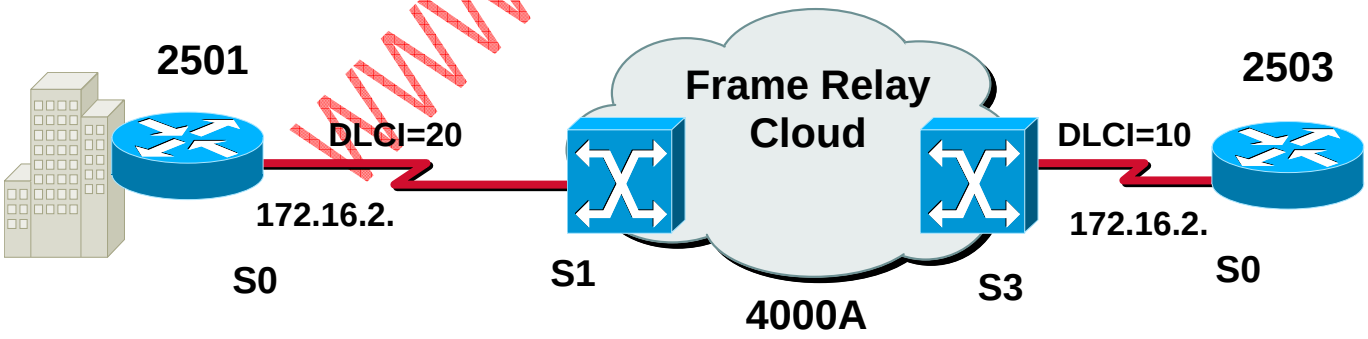
映射 VC 的目标与源的方式：

- 静态，Frame-map(帧中继映射)：远端 IP 与本地的 DLCI 的对应关系。
- 动态，inverse-arp(反转 arp)：慢且易出错。是缺省值。



RC 的帧中继配置：

```
Int s0
  Encap frame ietf                          (设置为 ietf 帧中继帧)
  No shutdown
  No frame inverse-arp                      (因为它慢且容易出错，所以改掉缺省值)
  Frame map ip 10.1.1.3   405 broadcast     (静态映射帧中继。帧中继缺省关闭广播，所以要打开广播)
  Frame map ip 10.1.1.2   105 broadcast     每条 VC 都映射。
```



帧中继交换机 CISCO 4000（DCE）

- **frame-relay switching**
- **interface Serial1**
- **bandwidth 64**
- **no ip address**
- **encapsulation frame-relay**
- **clockrate 2000000**
- **frame-relay lmi-type cisco**
- **frame-relay intf-type dce**
- **frame-relay route 200 interface Serial3 100**
- **interface Serial3**

- bandwidth 64
- no ip address
- encapsulation frame-relay
- no ip mroute-cache
- clockrate 2000000
- frame-relay lmi-type cisco
- frame-relay intf-type dce
- frame-relay route 100 interface Serial1 200

CISCO 2503（DTE）

- interface Serial0
- ip address 172.16.2.1 255.255.255.0
- encapsulation frame-relay
- frame-relay interface-dlci 100
- frame-relay lmi-type cisco
- frame-relay map ip 172.16.2.2 100 broadcast

CISCO 2501 （DTE）

- interface Serial0
- ip address 172.16.2.2 255.255.255.0
- encapsulation frame-relay
- frame-relay interface-dlci 200
- frame-relay lmi-type cisco
- frame-relay map ip 172.16.2.1 200 broadcast

Frame Relay 调试命令

- Show frame pvc
- 显示经过路由器的所有 PVC 的状态
- Show frame lmi
- 显示本地管理接口，LMI 为 VC 提供状态管理和广播；
- Show frame route
- 显示帧路由信息；
- Show frame map
- 查看当前映射项和 DLCI 映射表的相关信息。
- Show interface
- 提供了相关路由器上的所有接口的信息，up ,up
- 注意：DCE 端的时钟频率，带宽，等等设置
- DTE 端的线路的 DLCI 号与帧中继的静态映射。

QoS:把局域网边缘路由器缓冲区中（因为园区网流量大于远程线路流量）的待发 IP 包分级，按级别先后发送。

- 只有在拥塞发生时，QoS 才能运行。
- 它不能解决拥塞，就像交通警察作用一样。只有增大远程线路的带宽才能解决拥塞。

IP 包的分级方法：（cisco 主要用这 2 种）

1. 基于 IP 包中的 COS 字段（占 3bit，值 0～7，7 最高），国际通用的。
   - 由边缘路由器的进入端来设置操作。（策略路由）
   - 由发送端主机设置
2. 用访问列表对 IP 包分级。

QoS 实施的重要的队列技术：

1. FIFO（first input first output）先入先出队列：简单，所以快，占用 CPU 时间少。
   当远程速率大于 T1=1.544Mb/s 时（认为相对是高速）缺省是 FIFO.
   一个会话为一个队列
   如同时有低流高敏（如 telnet），高流低敏（如 ftp）的会话同时应用，fifo 就不合适了。
2. WFQ（加权公平队列）：
   当远程速率小于 T1=1.544Mb/s 时（认为相对是高速）缺省是 WFQ
   多队列存在
   基于会话来分队列的，一个会话一个队列。
   先发低流高敏的，余下高流低敏的会话队列，平均利用带宽（相同的时间片，即时分复用）
   开启 WFQ 的指令：int s0
   fair-qvere
3. CBWFQ（class base WFQ）: 最实用的
   可手动对会话队列分类
   可对每种队列设置精确带宽（一个访问列表为一个队列）

用 CBWFQ 配置：10.1.1.0/24 占 30%，10.1.2.0/24 占 30%，10.1.3.0/24 占 20%的远程带宽。

1. 把这 3 类包挑出来：

```
access-list 1 permit 10.1.1.0   0.0.0.255
access-list 2 permit 10.1.2.0   0.0.0.255
access-list 3 permit 10.1.3.0   0.0.0.255
```

2. 定义对应 IP 包的队列：

```
class-map 1                     (1 是队列标识符，也可以是字母)
    match ip address 1          (匹配列表 1)
class-map 2                     (2 是队列标识符，也可以是字母)
    match ip address 2          (匹配列表 2)
class-map 3                     (3 是队列标识符，也可以是字母)
    match ip address 3          (匹配列表 3)
class-map defult                (不匹配这 3 个队列都进入 defult 队列)
```

3. 赋予每个队列带宽属性：

```
policy-map lalala               (lalala 是带宽分配方案名)
    class 1
      bandwidth 30 percent      (带宽占 30%)
    class 2
      bandwidth 30 percent      (带宽占 30%)
    class 3
      bandwidth 30 percent      (带宽占 20%，剩下 20%会自动分配给 defult 队列)
```

4. 调用带宽分配方案

```
int s0
    service policy lalala out
```

LLQ 队列：绝对优先队列（江泽民），LLQ 队列必须传完了（其缓冲区 LLQ 传完），才能传其它队列。

其命令格式：如 class 1

priority (队列 1 为 LLQ 队列)

尾部丢包：缓冲区已被填满，某队列尾部被丢失。导致这些数据会被发送者重发。

Wred 技术（主动丢包）：当缓冲区占用到 80%时，优先级低（根据 cos 字段）的包会被扔掉，让高优先级的包，能够进入缓冲区。

其命令格式：如 class 1

bandwidth 30 percent

random-detect (随机检测，应用 wred 技术主动丢包)

VPN:
目标：低费用：DDN 是每月约 6000 元/月，帧中继是每月约 3000 元/月，用低廉的远程连接几个园区网。
安全：
相关的这些技术统称：VPN

1. 保证数据的机密性（让别人看不懂），方法：加密，暗文
2. 保证数据的完整性（让别人不能修改）
3. 源起证明。（身份认证）

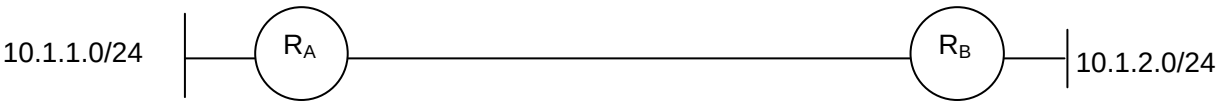
1. 机密性：（对 IP 包的每个字节的加密）
   a) 对称密码系统（pre-share）：加、解密的 KEY 是一样的（这样的快一些）
      i. des: key 为 56 位（70 年代的技术）
      ii. aes: key 为 128 位（测试阶段，还未实用）
      iii. 3des: key 为 3 个 56 位。（现在用的）
   b) 公匙、私匙：（慢）
      i. 公匙与私匙唯一对应（就好像存折的存折号和存折的密码的关系）
      ii. RAS 标准
2. 完整性：
   用散列算法实现保证完整性。一个任意长的数据，用散列函数可算出 64 字节的数据标签来对应它。标签放在包里，收者对比是否一致。
   散列算法标准：SHA
   HMA ：用 key 加密
   MD5：不用 key 加密

Ipsec(IP security)技术：

1. AH：认证 IP 头（头是路由信息），其中 TTL 字段、FCS 校验和不能加入。
   可用于身份证明、完整性 2 个方面。
2. ESP：对 IP 包的数据部分处理。
   身份证明、完整性、机密性 3 个方面都可引用。

SA(安全结合者)：相连 2 个路由器，IPsec 的参数协商统一后的结果。
一般会放在本地的 SA 的数据库中。
Ike(internet key 交换)：是一组协议。
目的：2 个路由器建立 SA
包括：iSAkmp(internet SA key management priority)等等协议

Ipsec 的指令配置：



1. 做访问列表，指定哪些包加密。

A:access-list 1 permit 10.1.1.0 0.0.0.255

2. 保证 Ike 的 SA，即保证协商的过程。

A:crypto isakmp policy 110 (110 是 Ike 策略的序号)
Auth pre-share (对称密码系统，保证第一次 Ike 的过程安全)
Encrypto des
Hash MD5 (散列标准：MD5)
Lefttime 6000 (110 只有 6000 秒的存活时间？)
Crypto iSAkmp key ciscokey ip 10.1.1.2 (ciscokey 是 key)

3. Ipsec 的 SA

Crypto Ipsec transfrom cisco esp-des AH-MD5 (cisco 是 Ipsec 的策略名，key 由 Ike 协商后决定的)

4. 用 map 组合以上步骤：

crypto map ciscomap 110 Ipsec-isakmp (Ipsec-isakmp 调用第二步)
match address 1 (这个 1 是指访问列表 1)
set peer 10.1.1.2
set transfrom ciscokey (ciscokey 对应第二步的那个 key)

5. 在接口上应用

int s0
crypto map ciscomap

B 上同样也要做这样的设定,只有第二、四步的那个 IP 变为 10.1.1.1

B: access-list 1 permit 10.1.1.0 0.0.0.255
crypto isakmp policy 110 (110 是 Ike 策略的序号)
Auth pre-share (对称密码系统，保证第一次 Ike 的过程安全)
Encrypto des
Hash MD5 (散列标准：MD5)
Lefttime 6000 (110 只有 6000 秒的存活时间？)
Crypto iSAkmp key cisco ip 10.1.1.1 (cisco 是 key)
Crypto Ipsec transfrom cisco esp-des AH-MD5 (cisco 是 Ipsec 的策略名，key 由 Ike 协商后决定的)
crypto map ciscomap 110 Ipsec-isakmp (Ipsec-isakmp 调用第二步)
match address 1 (这个 1 是指访问列表 1)
set peer 10.1.1.1
set transfrom ciscokey (ciscokey 对应第二步的那个 key)
int s0
crypto map ciscomap

比较 Internet 连接方案

NAT 与路由器比较
（NAT 安全性比较好，但是占用 CPU 处理时间比较多，不支持 IPSEC）

NAT 与代理服务器比较
（二者都能限制对内部网络的访问，都提供地址转换功能；代理服务器使用 TCP/UDP 端口转发信息包，这使得代理服务器能够执行更好的安全检查；代理服务器还支持高速缓存，可提高更好的速度)

Internet 连接共享与 NAT 比较

武汉热线上网服务器 IP: 202.103.24.116
武汉热线 DNS: 202.103.24.68

内部本地地址(Inside local IP Address):私有 IP，不能直接用于互连网。
内部全局地址(Inside Global IP Address)：用来代替内部本地 IP 地址的，对外，或在互联网上是合法的的 IP 地址。

NAT 功能:
内部网络地址转换
复用内部的全局地址
TCP 负载均衡
解决网络地址重叠

NAT 有三种类型：静态 NAT（staticNAT）、NAT 池（pooledNAT）和端口 NAT（PAT）。
1. 其中静态 NAT 设置起来最为简单，内部网络中的每个主机都被永久映射成 外部网络中的某个合法的地址。多用于服务器。
2. 而 NAT 池则是在外部网络中定义了一系列的合法地址，采用动态分配的方法映射到内部网络。多用于网络中的工作站。
3. PAT 则是把内部地址映射到外部网络的一个 IP 地址的不同端口上。

静态 NAT 配置：
把内部本地地址映射到内部全局地址（Maps the inside local address to the inside global address）.

```
interface Ethernet0
ip address 172.16.1.1 255.255.255.0
ip nat inside                                   （指定内部接口）
!
interface Serial0
 ip address 200.1.1.1 255.255.255.0
 ip nat outside                                 （指定外部接口）
!
ip nat inside source static 172.16.1.3 200.1.1.1   (建立两个 IP 地址之间的静态映射)
ip classless
ip route 0.0.0.0 0.0.0.0 200.1.1.2
```

注意：
从外网到内网建立静态映射后，外网能 PING 通内部全局地址（200.1.1.1）,如果使用真实地址，则访问失败，这是因为从外网没有到达内网的路由存在！

```
Ping 172.16.1.1 ……
Ping 200.1.1.1   !!!!!
```

动态 NAT 配置（NAT 池）
解析从内部本地地址 10.1.1.0/24 到内部全局地址 192.168.2.0/24 的唯一对应（Translate between inside hosts

addressed from 10.1.1.0/24 to the globally unique 192.168.2.0/24 network）

```
ip nat pool dyn-nat 192.168.2.1 192.168.2.254    netmask 255.255.255.0
ip nat inside source list 1 pool dyn-nat
!
interface Ethernet0
 ip address 10.1.1.10 255.255.255.0
 ip nat inside                                   （指定内部接口）
!
interface Serial0
 ip address 172.16.2.1 255.255.255.0
 ip nat outside                                  （指定外部接口）
!
access-list 1 permit 10.1.1.0 0.0.0.255
!
```

Configuring Inside Global Address Overloading

```
ip nat pool ovrld-nat 192.168.2.1 192.168.2.2    netmask 255.255.255.0
ip nat inside source list 1 pool ovrld-nat overload
!
interface Ethernet0/0
 ip address 10.1.1.10 255.255.255.0
 ip nat inside
!
interface Serial0/0
 ip address 172.16.2.1 255.255.255.0
 ip nat outside
!
access-list 1 permit 10.1.1.0 0.0.0.255
```

Basic IP address translation

```
Router#show ip nat trans
    Pro      Inside global   Inside local   Outside local   Outside global
    ---      192.2.2.1       10.1.1.1       ---             ---
    ---      192.2.2.2       10.1.1.2       ---             ---
```

IP address translation with overloading(A translation for a Telnet is still active. Two different inside hosts appear on the outside with a single IP address)

```
Router#sh ip nat trans
    Pro   Inside global        Inside local         Outside local        Outside global
    tcp   192.168.2.1:11003    10.1.1.1:11003       172.16.2.2:23        172.16.2.2:23
    tcp   192.168.2.1:1067     10.1.1.1:1067        172.16.2.3:23        172.16.2.3:23
```

Unique TCP port numbers are used to distinguish between hosts.